Paratuberculosis in dairy cattle

epidemiological studies used for design of a control programme in Denmark

Nielsen, Søren Saxmose

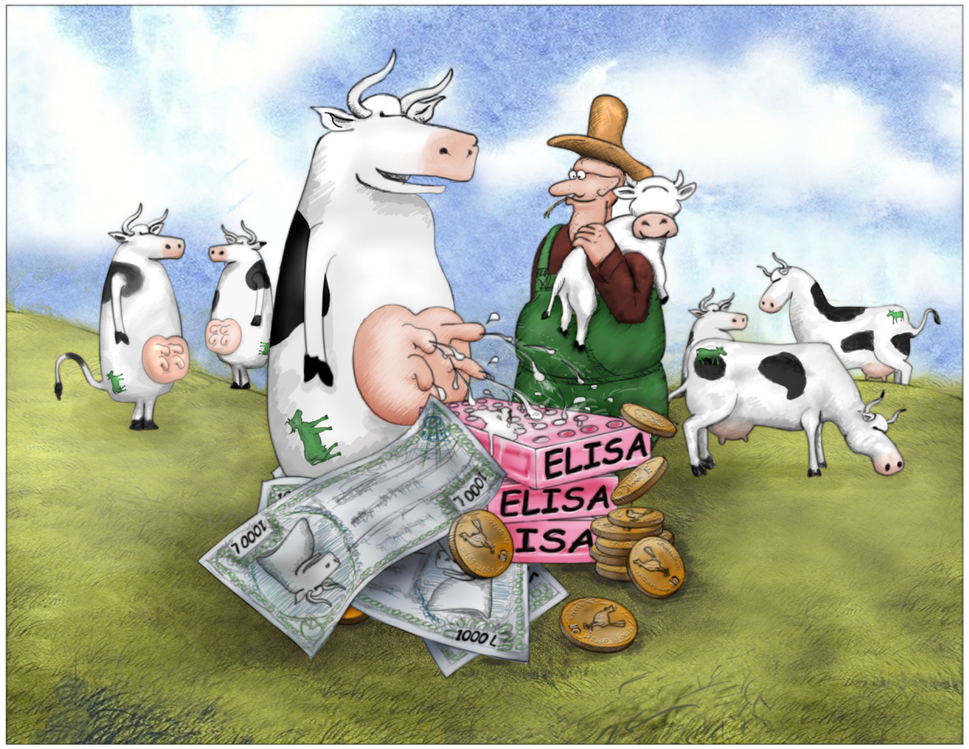
Publication date:
2009

Document version
Publisher's PDF, also known as Version of record

DEPARTMENT OF LARGE ANIMAL SCIENCES
FACULTY OF LIFE SCIENCES
UNIVERSITY OF COPENHAGEN

Paratuberculosis in Dairy Cattle

– Epidemiological studies used for design of a control programme in Denmark



DR.MED.VET. THESIS · SØREN SAXMOSE NIELSEN · 2009

Paratuberculosis in Dairy Cattle · Søren Saxmose Nielsen · Dr.med.vet. thesis · 2009

ISBN 978-87-7611-320-9

SL grafik

Paratuberculosis in Dairy Cattle

– Epidemiological studies used for design of a control programme in Denmark

by

Søren Saxmose Nielsen

Department of Large Animal Sciences
Faculty of Life Sciences
University of Copenhagen
Grønnegårdsvej 8
DK-1870 Frederiksberg C
Denmark

Denne afhandling er af den faglige referencegruppe for Husdyrbrugs- og Veterinærvidenskab under Akademisk Råd ved Det Biovidenskabelige Fakultet for Fødevarer, Veterinærmedicin og Naturressourcer, Københavns Universitet antaget til offentligt forsvar for den veterinærvidenskabelige doktorgrad

København, den 9. oktober 2009

Merete Fredholm
Ordførende for den faglige referencegruppe
for Husdyrbrugs- og Veterinærvidenskab under Akademisk Råd

Paratuberculosis in Dairy Cattle
- Epidemiological studies used for design of a control programme in Denmark

ISBN 978-87-7611-320-9

Front page artwork by Tim Evison, www.scientificillustration.net

Printed by: SL grafik, Frederiksberg C, Denmark (www.slgrafik.dk)

Table of Contents

This dissertation is based on the following 13 original articles, which are referred to in the text by their roman numerals (I-XIII):

I. Nielsen SS, Toft N, 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques. Vet Microbiol. 129: 217–235.
II. Nielsen SS, Ersbøll AK, 2006. Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows. J Dairy Sci. 89: 4557–4566.
III. Toft N, Nielsen SS, Jørgensen E, 2005. Continuous-data diagnostic tests for paratuberculosis as a multi-stage disease. J Dairy Sci. 88: 3923–3931.
IV. Nielsen SS, Toft N. 2006. Age-specific characteristics of ELISA and fecal culture for purpose specific testing for paratuberculosis. J Dairy Sci. 89: 569–579.
V. Nielsen SS, 2008. Transitions in diagnostic tests used for detection of *Mycobacterium avium* subsp. *paratuberculosis* infections in cattle. Vet Microbiol. 132: 274–282.
VI. Nielsen SS, Krogh MA, Enevoldsen C, 2009. Time to occurrence of drop in milk production in cows with various paratuberculosis antibody profiles. J Dairy Sci. 92: 149–155.
VII. Nielsen SS, Toft N, 2009. A review of prevalences of paratuberculosis in farmed animals in Europe. Prev Vet Med. 88: 1–19.
VIII. Sergeant ESG, Nielsen SS, Toft N, 2008. Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds. Prev Vet Med. 85: 92–106.
IX. Nielsen SS, Toft N, Jørgensen E, Bibby BM, 2007. Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response. Prev Vet Med. 81: 290–305.
X. Nielsen SS, Gröhn YT, Quaas RL, Agger JF, 2002. Paratuberculosis in dairy cattle: Variation of the antibody-response in offspring attributable to the dam. J Dairy Sci. 85: 406–412.
XI. Nielsen SS, Bjerre H, Toft N, 2008. Colostrum and milk as risk factors for infection with *Mycobacterium avium* subsp. *paratuberculosis* in dairy cattle. J Dairy Sci. 91: 4610–4615.
XII. Nielsen SS, Toft N, 2007. Assessment of management-related risk factors for paratuberculosis in Danish dairy herds using Bayesian mixture models. Prev Vet Med. 81: 306–317.
XIII. Kudahl AB, Nielsen SS, Østergaard S, 2008. Economy, efficacy and feasibility risk-based control program against paratuberculosis. J Dairy Sci. 91: 4599–4609.

ACKNOWLEDGEMENTS

The research presented in this thesis was performed at the Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Denmark (formerly The Royal Veterinary and Agricultural University) and was supported by the following funding bodies:

- Danish Cattle Levy Fund
- Danish Milk Levy Fund
- The European Commission, 6th Framework Programme, project no. 023106 (FOOD) (ParaTBTools)
- Danish Directorate of Food, Fisheries and Agro Business (DFFE, J.nr. 3412-05-1092 and J.nr. 3412-06-01550)
- The Danish Cattle Federation

I am grateful for all the financial support I received. Furthermore, I am profoundly grateful to the Danish Dairy Board and the Danish Cattle Federation for providing a framework, which could be used for many of the studies. The cooperation with the dairy industry has been truly remarkable. I am very appreciative of the assistance and discussions with employees from the Danish Dairy Board.

I would like to express my sincere thanks to all my co-authors and colleagues. I have been blessed with extremely intelligent, critical and helpful friends, whom have contributed greatly to the work carried out. Many deserve to be mentioned by name, but I hope you are all aware of my appreciation without listing you all. Some have helped a little, some a lot. All help has made a huge difference to me.

Lastly, my gratitude goes to my family whom have continuously supported my work and my dreams. No critical questions, only enthusiasm and pride of my achievements. Thank you all, especially to Arshnee, who has suffered many a time by hearing or reading the phrase “Paratuberculosis is a chronic infection…….” Arshnee, I owe you so much. Thank you!

Frederiksberg, February 2009
Søren Saxmose Nielsen

PREFACE

Paratuberculosis is a chronic infection of the intestine caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). The infection affects cattle and other ruminants. The Danish Dairy Board in the late 1990s concluded that adverse effects of MAP infections should be investigated further, and in 1999 several projects on paratuberculosis were initiated. These projects were part of a larger project called Kongeåprojektet (Andersen et al., 2000). Several other studies of paratuberculosis were subsequently funded by the cattle industry and public institutions.

The fact that MAP infections are *chronic* suggests that time should be an essential component in studies of the infections. Still, important aspects of the infection such as the incubation period remains poorly characterised. Precisely because the infection is chronic its effects vary from one infection stage to another, yet most studies of the pathogenesis, diagnosis and effects of the disease have been of cross-sectional designs. This study design may not be ideal for studies of a chronic infection. The effects of MPA infections are multiple, e.g. infection, reduced milk production and transmission of MAP to other animals, and death. These effects can occur at multiple levels e.g. cow and herd-level. Differences in effects and level of decision making, warrant different purposes of testing. There was therefore a need for longitudinal studies characterising diagnostic test performance in relation to purpose of testing and the effects that MAP infections cause. A finding that the effects were significant would allow decision makers to examine the prevalence of MAP infection and draw inferences about control of the infection. Therefore, unbiased estimates of MAP infection are required for decision-making, as well as for use in risk factor studies.

In 2005 farmers requested that a control programme on paratuberculosis should be established. “Operation Paratuberculosis” was launched in February 2006, with the aim of providing tools for on-farm management of paratuberculosis as well as reducing its national prevalence. Operation Paratuberculosis was to a large extent founded, scientifically speaking, on the results presented here and elsewhere in literature.

This thesis aims to provide an overview of the epidemiological studies conducted in Denmark and used in decision processes leading to the design of the Danish control programme on paratuberculosis in dairy cattle. Sociological and political issues also need to be addressed in the establishment of a control programme, but these issues are not discussed in this thesis

SUMMARY

Mycobacterium avium subsp. *paratuberculosis* (MAP) causes chronic infections in dairy cattle worldwide, resulting in economic losses to dairy farmers. This infection has existed for many years. Its control is hampered by inaccurate diagnostic tests. These tests are not very sensitive tools for detecting MAP-infected animals due to the chronic nature of the infection. MAP infections can be insidious, and economic losses may only be realised after the bacteria have been present in a herd for many years.

The aim of this thesis was to describe a cost-effective, risk-based approach to control of MAP infections in dairy herds. Specifically, epidemiological studies were conducted to a) characterise the performance of diagnostic tests; b) establish prevalences; and c) identify risk factors. The information was then compiled in the description and evaluation of a control scheme. This work is described in seven chapters, which are supported by studies described in thirteen accompanying papers.

The background of the thesis was: Diagnosis of MAP infections can be obscured by low specificity and low sensitivity, resulting, respectively, in false-positive and false-negative test results. MAP belongs to *M. avium*, which has four subspecies and is widespread in the environment. Non-MAP *M. avium* infections may result in false-positives in tests used to detect MAP infections, reducing test specificity.

The sensitivity of most tests is affected by the chronic nature of the infection. Infections can be latent for an unknown period of time. Relatively little information is available on infection dynamics and the diagnostic test responses in the course of a MAP infection, because only a few longitudinal studies have been carried out. The exact timing of infection is unknown, although MAP infections are primarily thought to occur in calfhood. In addition, it is unclear whether all infections result in adverse effects, such as bacterial shedding, reduced milk yield, weight loss and eventual death. Nor is it clear whether pro-inflammatory immune reactions clear the infection or manage to keep it under control. Because end-stage disease can occur at any time during an animal's life, the incubation period appears to be variable.

A review of the accuracies of diagnostic tests revealed large variation in the performance of various ELISAs and faecal culture. However, results of the present thesis showed that distinction between different stages of infection involving the categories "MAP-infected", "MAP-infectious" and "MAP-affected" increased the ease of interpretation. The diagnostic tests are not particularly sensitive in detecting infected animals. However, immunity-based diagnosis relative to adverse effects may be a useful way to control MAP infections, because the occurrence of humoural immune responses seems to correspond with deterioration of the infection. If the tests are used in relation to adverse effects, they should be evaluated in relation to each purpose. The studies in the review were all cross-sectional, but an indirect antibody ELISA for detection of MAP-infected and MAP-infectious animals in a longitudinal study design was characterised in this thesis.

The results suggested that sensitivity in the detection of MAP-infected animals increased with age, whereas the probability of detecting MAP infectious animals was less affected by the age of the tested animal. It was also demonstrated that most cows shedding MAP would test positive in the antibody ELISA at some point in time. Furthermore, it was demonstrated that antibodies to MAP occurred 3–4 years prior to the detection of MAP shedding by faecal culture in approximately 20% of cows classified as high shedders. Antibodies may therefore occur relatively early in some infected animals. The results also suggested that testing animals just once, one year before detected shedding would result in the detection of only 30% of the animals that became high shedders. Testing in the interval 0–3 months prior to onset of detectable shedding would result in detection of up to 60–70% of the high shedders. The longitudinal studies also suggested that in general ELISA may result in positive tests prior to onset of detectable bacterial shedding. Higher ELISA values are equivalent to a higher probability of MAP shedding. Finally, it was demonstrated that the milk yield of animals with fluctuating antibody reactions can be higher than that of repeated test-negative herd-mates. Animals with the last milk sample being positive or repeated positive ELISA-tests can display a significant drop in milk production. Therefore, frequent testing and ordinal ELISA values can be used to detect animals with a high risk of MAP shedding and, in combination with milk production data, can be used to determine which animals should be considered for culling.

ELISA is not a very sensitive detector of MAP-infected animals. Yet it has been the test most widely used to estimate prevalences. A review of prevalences in Europe suggested that most prevalence estimates are not comparable, partly as a result of the variable test sensitivity estimates. Therefore, it is difficult to estimate within-herd prevalences and certify herds free of infection. A Bayesian mixture model was developed. It was found to be more precise than a cut-off based model. The mixture model was subsequently used to assess management factors associated with the development of antibodies to MAP.

To control MAP infections, information about transmission is needed. The transmission of MAP is thought primarily to occur via the faecal-oral route, but MAP can also be transmitted via infected milk or *in utero*. A retrospective and a prospective study were conducted to assess risk factors for transmission of MAP. It was demonstrated that the ELISA status of a dam influenced the ELISA status of her calf. Calves fed colostrum from multiple cows rather than just one cow had a higher risk of having MAP antibodies as an adult. In addition, the “source of milk” only appeared as an important risk factor when calves were suckling milk from a foster cow. Housing type and animal density of calves and young stock appeared to affect the risk of developing antibodies. The identified risk factors along with conceptual and previously identified risk factors suggest that efforts to control MAP should include measures to ensure that the manure from adults is kept from susceptible animals (mainly calves). Furthermore, milk should be considered infectious, and *in utero* infections are likely to occur.

Specific management practices can be adopted to deal with the risks mentioned above (e.g. calves can be removed from their dam immediately after birth). However, in many farming production systems these practices are considered time-consuming and therefore will not be implemented. A risk-based system that is tailored to each farm may be an acceptable option to herd managers.

A risk-based system with identification of high- and low-risk animals was investigated as a cost-effective solution requiring fewer animals to be managed. Testing with milk ELISA would be conducted four times per year to identify high-risk animals. Special management procedures to avoid transmission from high-risk animals should be established; low-risk animals can be given less attention. High-risk animals thought to contribute to high bacterial loads in the environment should be culled prior to next calving. Simulation studies indicated that in production systems, where a high workload is required to reduce the risk of transmission of MAP at calving, a risk-based approach would be cost-effective. A non-risk-based approach would be more cost-effective in systems where all animals are managed with high bio-security standards at low cost. A drawback of the latter would be that no data are available to monitor changes in MAP prevalence.

There are still insufficient data to show proof of concept, but frequent testing has previously proven to be a useful strategy for a chronic insidious infection – the infection being tuberculosis caused by *M. bovis* in cattle. Certification of freedom from MAP in a herd is not yet possible owing to the low sensitivities of tests detecting MAP infections, but continued monitoring with ELISA may provide historical data which, in a Bayesian framework, can be used to increase the estimated probability that a herd is free of MAP.

The results in this thesis demonstrate that, used appropriately, milk ELISA can be a cost-effective method for controlling MAP infections in dairy cattle. Frequent testing is necessary, along with changes in management designed to reduce transmission. The role, in MAP transmission, of further management factors should be explored in order to optimise control schemes.

SAMMMENDRAG

Mycobacterium avium subsp. *paratuberculosis* (MAP) er årsag til kroniske infektioner hos malkekvæg og giver anledning til økonomiske tab i kvægindustrien. Infektionen har eksisteret i mange år, men sanering vanskeliggøres af mangel på gode diagnostiske tests. Testene har lav sensitivitet i forbindelse med påvisning af MAP-inficerede dyr som følge af infektionens kroniske natur, specielt i de tidligere stadier af infektionen. MAP infektioner kan være latente, og de økonomiske tab vil ofte først være synlige, når bakterierne har været tilstede i en besætning i mange år.

Formålet med denne afhandling var at beskrive en omkostningseffektiv risikobaseret tilgang til sanering for MAP i malkekvægsbesætninger. Mere specifikt blev epidemiologiske studier gennemført for at beskrive og karakterisere a) sensitivitet og specificitet af diagnostiske test; b) prævalensen af MAP infektioner; og c) risikofaktorer for smittespredning med MAP. Denne information blev efterfølgende sammenfattet i beskrivelsen og evalueringen af et saneringsprogram. Arbejdet er beskrevet i syv kapitler, som understøttes af studier beskrevet i tretten medfølgende artikler.

Diagnose af MAP infektioner kan vanskeliggøres af lav sensitivitet og specificitet, som giver henholdsvis falskpositive og falsknegative testresultater. MAP tilhører *M. avium* arten, som er hyppigt forekommende i miljøet, og som har fire underarter. *M. avium* infektioner, som ikke skyldes MAP, kan resultere i falskpositive resultater, hvilket reducerer testspecificiteten af MAP test.

Sensitiviteten af de fleste test er i høj grad påvirket af sygdommens kroniske natur, idet infektionerne kan være latente igennem længere tid. Der eksisterer relativt begrænset med information om infektionsdynamikken og de diagnostiske testresultater i forløbet af infektionen, fordi der kun har været lavet få longitudinelle studier. Infektionstidspunktet er ukendt, selvom de fleste MAP infektioner menes at blive etableret i kalvestadiet. Desuden er det uklart, om alle infektioner resulterer i negative effekter såsom bakterieudskillelse, reduceret mælkeydelse, vægttab og død. Det er heller ikke klart om proinflammatoriske immunreaktioner kan eliminere infektionen eller blot holde den på et niveau, hvor den er under kontrol. Inkubationsperioden lader til at variere, da udbrud af sygdom kan ske på et hvilket som helst tidspunkt i et dyrs liv.

Et review af sensitivitet og specificitet af diagnostiske test afslørede stor variation for både ELISA og bakteriologisk dyrkning. Stratificering i forhold til sygdomsstadierne ”MAP-inficeret”, ”MAP-infektiøs” og ”MAP-afficeret” kunne dog lette fortolkningen af testresultaterne. De diagnostiske test er ikke specielt velegnede til at påvise MAP-inficerede dyr, men test til påvisning af antistoffer kan være brugbare til sanering for MAP infektioner, da påvisning af humorale immunreaktioner er korreleret med sygdomsudviklingen. Hvis de diagnostiske test bruges i relation til de uønskede effekter, bør de således evalueres i relation til de specifikke formål. Studierne i oversigtsartiklen var alle tværsnitsstudier, men en

indirekte antistof ELISA til påvisning af MAP-inficerede og MAP-infektiøse dyr blev karakteriseret i et studium med et longitudinelt design i denne afhandling.

Resultaterne viste, at sensitiviteten til påvisning af MAP-inficerede dyr blev højere med stigende alder, mens sandsynligheden for at detektere MAP-infektiøse dyr var mindre påvirket af alderen af det testede dyr. Det blev også vist, at de fleste dyr, der udskiller MAP, på et tidspunkt vil teste positive i antistof ELISA. Det blev endvidere demonstreret, at antistoffer mod MAP kunne påvises 3–4 år før påviselig udskillelse af bakterier i ca. 20% af de køer, som senere blev klassificeret som høj-udskillere. Antistoffer kan derfor optræde relativt tidligt i nogle inficerede dyr. Resultaterne viste også, at hvis et dyr blev testet et år før det blev høj-udskiller, ville kun ca. 30% af dyrene være antistofpositive. Blev dyret derimod testet i intervallet 0–3 måneder før bakterieudskillelsen begyndte, ville 60–70% af høj-udskillerne blive testpositive før bakterieudskillelsen begyndte. De longitudinelle studier antydede endvidere, at ELISA generelt vil være positiv før påviselig bakterieudskillelse forekommer. Slutteligt blev det demonstreret, at mælkeydelsen hos dyr med svingende antistofniveauer kan være højere end hos dyr, der har været testnegative gentagne gange. Dyr hvis seneste ELISA-resultater var positive kan have et væsentligt fald i mælkeydelsen. Hyppige test og brug af ELISA-resultater på en ordinal skala kan bruges til at påvise dyr, som har en høj risiko for udskillelse af bakterier, og i kombination med mælkeproduktionsdata kan de bruges til at tage beslutninger om hvilke dyr, som bør udsættes.

ELISA er ikke specielt følsom til påvisning af MAP-inficerede dyr, men er alligevel den hyppigst benyttede test til at estimere prævalenser. En gennemgang af prævalenserne i Europa viste, at prævalensestimaterne ofte ikke er sammenlignelige, delvist som følge af meget variable og upræcises testsensitivitetsestimater. Det er derfor også vanskeligt at estimere prævalensen indenfor besætninger og certificere besætningerne fri for MAP infektioner. Hyppigt har man anvendt cut-off baserede tilgange til at opdele test-resultaterne i positive og negative, men i stedet blev en bayesiansk mixturmodel udviklet. Den blev fundet at være mere præcis end den cut-off-baserede metode, og mixturmodellen blev efterfølgende brugt til at evaluere hvilke managementfaktorer, som er associeret med udvikling af antistoffer mod MAP.

Information om smittespredning er nødvendig for at kontrollere MAP infektioner. Smittespredning menes hovedsagligt at forekomme via oral optagelse af bakterier, som udskilles i fæces, men MAP kan også udskilles i mælk og overføres *in utero*. Et retrospektivt og et prospektivt studium blev udført for at evaluere risikofaktorer for smittespredning med MAP. Det blev vist, at en mors antistofstatus var associeret med antistofstatus hos hendes kalv, og kalve, der blev fodret med råmælk fra mange køer frem for kun en ko, havde en højere risiko for at udvikle antistoffer mod MAP. Kalve, som gik med ammekøer, havde også en højere risiko for at udvikle antistoffer. De fundne risikofaktorer sammen med hypotetiske og tidligere identificerede risikofaktorer peger på, at sanering for MAP skal inkludere tiltag,

som sikrer modtagelige dyr (hovedsagligt kalve) ikke kommer i kontakt med gødning fra voksne dyr. Yderligere skal mælk opfattes som en mulig smittekilde, og *in utero* infektioner forekommer hos dyr født af mødre med antistoffer mod MAP.

Specifikke tiltag kan implementeres i en besætning for at håndtere de risici, som er nævnt ovenfor (f.eks. kan kalve fjernes fra deres mor lige efter fødslen). Nogle tiltag vil dog i visse besætninger blive opfattet som meget tidskrævende og vil derfor ikke blive implementeret. Derfor kan et risikobaseret system, som tilpasses den enkelte besætning, være en acceptabel løsning for mange driftsledere. Et risikobaseret system baseret på identifikation af høj- og lav-risiko dyr blev undersøgt som en omkostningseffektiv løsning, hvor færre dyr skal håndteres i dagligdagen. Udpegning af risikodyrene blev baseret på fire årlige besætningsscreeninger med en antistof-ELISA udført på mælkeprøver. Driftslederen skal etablere managementprocedurer, så smittespredning fra høj-risikodyr undgås, mens disse procedurer ikke er nødvendige for lav-risikodyrene. De høj-risikodyr, som menes at bidrage med størst bakterieudskillelse, skal udsættes før næste kælvning. Det blev med simuleringsstudier vist, at i de produktionssystemer, hvor en stor arbejdsindsats er påkrævet for at reducere risikoen for smittespredning i forbindelse med kælvning, vil en risikobaseret tilgang være den mest omkostningseffektive. En ikke-risikobaseret tilgang kan bruges i systemer, hvor et højt hygiejneniveau kan etableres til lave omkostninger, men ulempen ved sidstnævnte er, at der ikke vil genereres data, som kan bruges til at monitorere udviklingen i prævalensen.

Der foreligger ikke data, som kan vise, at konceptet virker i den virkelige verden, men hyppig testning har tidligere været benyttet med succes til en kronisk, snigende infektion, nemlig ved kvægtuberkulose forårsaget af *M. bovis*. Certificering af besætninger som fri for MAP er endnu ikke mulig som følge af den lave sensitivitet af tilgængelige diagnostiske tests, men fortløbende monitorering af besætningerne med antistof ELISA kan bidrage med historiske data, som med en bayesiansk tilgang kan bruges til at estimere sandsynligheden for, at en besætning er fri for MAP.

Resultaterne i denne afhandling viser, at hvis antistof ELISA på mælk anvendes korrekt, kan testen bruges som led i omkostningseffektive saneringer for MAP infektioner hos malkekvæg. Hyppige test er nødvendige i kombination med ændringer i management således, at smittespredning reduceres. Andre managementfaktorer af betydning for smittespredning kan med fordel undersøges for at optimere saneringsindsatsen yderligere, og de tilgange, som er demonstreret her, kan bruges som skabeloner i fremtidige studier heraf.

INTRODUCTION

Mycobacterium avium subsp. *paratuberculosis* (MAP) is the causative agent of a chronic bacterial infection of the intestine affecting cattle and other ruminants. MAP infections are costly to the cattle industry, with estimated losses of 250 million USD / year in the US dairy industry (Ott et al., 1999). There are limited data on economic losses for other countries, but many major dairy production countries either already run, or intend to implement, MAP control, surveillance or eradication programmes (Kennedy and Nielsen, 2007). These programmes have been established mainly to limit both short-term adverse effects, such as reduced milk yield, premature culling, weight loss and death, and long-term effects, which include continued transmission of MAP and loss of genetic potential. Food safety concerns – e.g. about MAP contamination in milk – can provide an additional reason for establishing a programme (Franken, 2005). Viable MAP have been isolated from commercially available pasteurised milk in the United Kingdom (Grant et al., 2002), and MAP is suspected of being involved in Crohn's disease in humans (e.g. Uzoigwe et al., 2007).

Despite these costs, some countries are reluctant to establish programmes because a) MAP is not considered a cause of large enough economic losses to justify the efforts (Kennedy and Nielsen, 2007); and b) traditional test-and-cull programmes are not sufficiently effective on their own and require changes in management to reduce transmission (Kudahl et al., 2007). Such changes may be difficult to implement (Ridge et al., 2005), because habits are hard to change, or because there are few inexpensive, cost-effective diagnostic tests.

A programme should therefore be cost-effective; its benefits should be obvious; changes in management should require limited additional workload; and information about the progression of control should be available. Limited changes to workload would be possible if only high-risk animals needed to be handled. Frequent diagnostic testing could be used to identify high-risk animals, and repeated testing could then be used to track potential adverse effects.

Hypothesis

The control of infections caused by MAP can be achieved cost-effectively through a risk-based control scheme in which high-risk cows are detected through repeated testing using a milk ELISA for the detection of MAP antibodies. Furthermore, control schemes can be improved by identifying management factors that reduce MAP transmission.

This working hypothesis was investigated through studies conducted in three main categories:

I) Diagnosis, where the accuracy of milk ELISA is characterised relative to different purposes of testing through longitudinal studies;

II) Prevalence estimation based on imperfect diagnostics, with discussion of alternatives to current approaches; and

III) Risk factor studies based on ranked prevalence estimates and with retrospective and prospective study designs.

The results were subsequently combined in the design of a MAP control programme.

Outline of the thesis

The research described in this thesis was carried out as a continuation of work presented in a doctor of veterinary medicine graduation thesis and a PhD thesis obtained from The Royal Veterinary and Agricultural University (KVL) in 1998 and 2002, respectively. The two theses were entitled: "*ELISA diagnosis of bovine paratuberculosis in Denmark*" (original Danish title: *ELISA-diagnostik af bovin paratuberkulose i Danmark*) and "*Paratuberculosis in Danish dairy cattle. Interpretation of diagnostic information depending on purpose and disease stage*". The modification and evaluation of an existing ELISA, along with a prevalence study, was described in the graduation thesis, and purposive interpretation of diagnostic test information was investigated in the PhD thesis, primarily on the basis of cross-sectional studies.

This thesis investigates the hypothesis mentioned above primarily by means of studies with longitudinal components. The studies are described in detail in the accompanying papers. A general introduction to MAP, and to the pathogenesis of, and effects caused by, MAP infections is described in Chapter 1, along with a discussion of the design of control programmes. The chapter focuses on factors of interest in the diagnosis and control of paratuberculosis. Chapters 2–4 cover the topics of diagnosis, prevalence estimation and risk factors, respectively. Each of these chapters introduces the topic, summarises current knowledge, and then reports and discusses the results described in the accompanying papers. Of the accompanying papers, I-VI concern diagnosis, VII-IX concern prevalence estimation, and X-XII concern risk factors. Chapter 5 deals with the design of a risk-based MAP control programme; it also provides technical and economic evaluation of the risk-based approach. In Chapter 6, conclusions drawn from the studies included in the thesis are summarised, and Chapter 7 presents future perspectives on work carried out on, and the control and eradication of, MAP infections in Denmark.

1. GENERAL INTRODUCTION TO PARATUBERCULOSIS

1.1. *Mycobacterium avium* subsp. *paratuberculosis*

MAP are slow-growing, acid-fast bacteria belonging to the species *M. avium*. They are non-spore forming, non-motile, rod-shaped obligate aerobes, which usually require the presence of mycobactin J, an iron-chelating agent, for growth *in vitro*. However, some other *M. avium* strains are also mycobactin-dependent. Visible growth usually takes 8–12 weeks at 37^{o}C (Hirsh and Biberstein, 2004). Three other *M. avium* subspecies have been associated with animals: *M. avium* subsp. *silvaticum* (MAS, previously called wood-pigeon mycobacteria) and *M. avium* subsp. *avium* (MAA, Thorel et al., 1990) and *M. avium* subsp. *hominissuis* (MAH, Mijs et al., 2002).

Molecular biology

Analysis of the MAP genome has identified a repetitive DNA sequence that is specific to MAP (Collins et al., 1989). This insertion sequence, IS*900*, has a size of 1.45 kbp, and the MAP genome includes 15 to 20 dispersed copies. IS*900* can be identified by PCR, and IS*900* PCR is therefore capable of distinguishing MAP from MAS, MAH and MAA, although the reference method for speciation of mycobacteria is 16S rDNA sequencing (Mijs et al., 2002).

Implications of non-MAP M. avium to MAP diagnosis

MAS, MAA and MAH are important in relation to MAP, because they and MAP appear to share both antigenic features that can cause false-positive reactions in immuno-diagnostic tests and phenotypic features, such as mycobactin-dependency, which sometimes yield false-positive results at culture. Furthermore, MAS, MAA and MAH may cause bovine disease which, in some cases, resembles paratuberculosis caused by MAP (Collins et al., 1983; 1985; Collins et al., 1997; Matthews and McDiarmid, 1979). In Denmark, MAA have frequently been isolated from cattle (Jørgensen, 1978). From the veterinary viewpoint, the effects caused by the *M. avium* subspecies are more important than their exact classification.

1.2. Transmission of MAP

MAP are shed in faeces resulting in contamination of the environment with subsequent uptake by susceptible animals (Sweeney, 1996). MAP can also be excreted in colostrum and milk (Taylor et al., 1981; Sweeney et al., 1992a; Streeter et al., 1995), which are often used for feeding of calves. MAP are most commonly ingested orally. Infections *in utero* have also been reported (Seitz et al., 1989; Sweeney et al., 1992b) and are estimated to occur in 9% of subclinically infected dams and 39% of clinically infected dams (Whittington and Windsor, 2009). MAP infection facilitated by *in vitro* and *in vivo* fertilisation with embryos from MAP-infected cows has been described as an unlikely event, even if MAP are present in material flushed from the donor cows (Bielanski et al., 2006). There are no data to suggest that MAP

can be transmitted via semen during artificial insemination, but the risk may be higher than negligible (Anon., 2004a).

Factors influencing transmission operate in five areas: a) source of infection, b) shedding of agent, c) mode of transmission, d) mechanism of invasion and e) host susceptibility (Houe et al., 2004). For MAP (Fig. 1.1), the source of infection is, by definition, MAP-infectious animals (Nielsen and Toft, 2008 I). MAP can be transferred directly, or they can be shed into the environment, where they can survive for a long time: survival of up to 252 days in slurry has been reported (Jørgensen, 1977). The environment may therefore also serve as a source of infection. Shedding of the agent occurs from MAP-infectious animals. The potential of diagnostic tests for detecting these infectious animals have been described in Chapter 2. MAP can be transmitted horizontally via oral uptake of milk and faeces containing MAP, and vertically via colostrum and *in utero* (Sweeney, 1996).

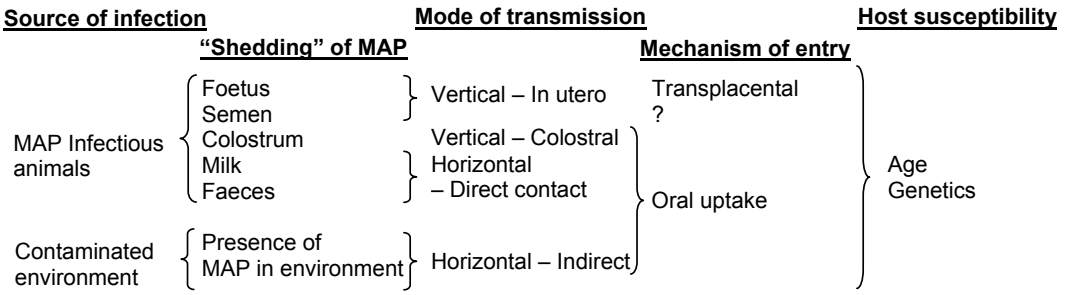


Fig.1.1. Schematic representation of factors influencing transmission of MAP

1.3. Pathogenesis

Susceptibility to MAP infection

MAP infections in cattle have been described as occurring primarily in calfhood (Taylor, 1953). Although resistance to infection appears to increase with age, infection may occur in adult cattle (Doyle, 1953) following high doses of MAP. Age-resistance may be less significant than is generally supposed, since data supporting the hypothesis of such resistance are few. A summary of the findings described in literature is given in Table 1.1. This summary suggests that MAP infection with or without detectable pathological lesions and clinical disease occurs at almost the same frequencies irrespective of age-group. However, the quality of the studies was generally poor, the infective doses were not comparable, and the necessary assurance that adult cattle were not naturally exposed to MAP in early life, prior to experimental infection, is generally missing. For example, Doyle (1953) imported adult animals from Ireland, where MAP prevalence was thought to be low (but not certified free) and used an infective dose of 60 g / animal. He scraped the infective material from the mucous membrane of the small intestine after it had been washed. Doyle managed to infect and produce clinical disease in 6 of his 12 adult study objects. Rankin

(1958; 1961a) used only 100 mg intravenously and produced no clinical disease in adult animals (0 of 5 adults) but did so in two calves (4 of 6 calves). However, Rankin (1961a) managed to infect the adult animals, as documented by detection of MAP, in 5 out of 5 inoculated adults 4 years after inoculation. The studies of Rankin (1958; 1961a) and Taylor (1953) appear to have a greater validity than the majority of other studies in Table 1.1. In these studies both young and old animals were included in the study groups, and the variation in outcome of the infection was small within study groups, whereas the between-group variation was large. This distribution of variation suggested that the infective doses used were sufficient to produce infection and disease, revealing potential differences in age-related susceptibility to infection. However, it is not certain that age-related resistance is as important as is generally believed; exposure dose and frequency are likely to affect the probability of an animal becoming infected (Hines et al., 2007).

Table 1.1. Summary of studies on the effect of age at exposure to MAP on pathogenesis in MAP infections

Age at infection	Mode of infection†	No. of animals					Reference
		In study	With pathology*	With MAP in intestinal or other tissues	Shedding in faeces	With clinical symptoms (median time PI)‡	
0 mo	IV (100 mg)	2	2	2	2	2 (400 DPI)	Taylor, 1953
0 mo	Oral (100 mg)	2	2	2	2	1 (400 DPI)	Taylor, 1953
1 week	Natural[d]	3	1	3	3	0 (culled 5 yrs PI)	Rankin, 1961b
1 mo.	IV (100 mg)	6	4	6	6	4 (9 mo.)	Rankin, 1958
1 mo.	Oral (180 mg)	2	2	2	2	0 (culled at 150 DPI)	Larsen et al., 1975
1 mo	Natural[d]	3	3	3	3	2 (2–3 yrs PI)	Rankin, 1961b
3 mo	IV (100 mg)	2	1	2	2	1 (400 DPI)	Taylor, 1953
3 mo	Oral (100 mg)	2	2	2	2	0 (culled 730–760 DPI)	Taylor, 1953
Calves	Natural[d]	23	13	NA	NA	7 (mean: 32 mo at death)	Hagan, 1938
Summary young calves		45	30/45 (66%)	22/22 (100%)	22/22 (100%)	17/45 (38%)	
6 mo	IV (3 mL)	1	1	1	NA	0 (culled 764 DPI)	M'Fadyean & Sheather, 1916
6 mo	Oral (6 mL)	1	1	1	NA	0 (culled 764 DPI)	M'Fadyean & Sheather, 1916
6 mo	IV (4 mL)	3	0	3	NA	0 (culled 440 DPI)	M'Fadyean & Sheather, 1916
6 mo	Oral (100 mg)	2	0	1	2	0 (culled 730–790 DPI)	Taylor, 1953
6 mo	IV (100 mg)	2	0	0	1	0 (culled 670–790 DPI)	Taylor, 1953
6 mo	Natural[d]	3	2	3	3	2 (3–4 yrs PI)	Rankin, 1961b
9 mo.	Oral (180 mg)	4	3	3	3	0 (culled at 150 DPI)	Larsen et al., 1975
<1 yrs[$]	Natural[d]	6	1	NA	NA	1 (37 mo at death)	Hagan, 1938
<1 yrs	Oral (60 g)	9	6	NA	NA	6 (median: ?)	Hagan, 1938
Summary older calves		31	14/31 (45%)	12/16 (75%)	9/11 (82%)	9/31 (29%)	
>1 yrs[$]	Natural[d]	6	0	NA	NA	0	Hagan, 1938
1–2 yrs	Oral (60 g)	3	3	NA	NA	3 (median: ?)	Hagan, 1938
Summary young stock		9	3/9 (33%)	NA	NA	3/9 (33%)	
>2 yrs	Oral (60 g)	3	0	NA	NA	0	Hagan, 1938
Adult (1.5–6 yrs)	Oral (60 g)	12	Min. 6	Min. 5	Min. 3	6 (17.5 mo PI)	Doyle, 1953
Adult (3 yrs)	IV (100 mg)	5	1	5	0	0 (culled 4–4.5 yrs PI)	Rankin, 1961a
4 adult (5–11 yrs)	Oral (180 mg)	4	2	2	0	0 (culled at 150 DPI)	Larsen et al., 1975
Summary adults		21	≥9/21 (≥43%)	≥12/21 (≥57%)	≥3/21 (≥14%)	6/21 (≥29%)	

*With pathology: Signs of thickened intestinal mucosa at gross pathological examination or histological lesions.
#With MAP in intestinal or other tissues: MAP had been cultured from one or more tissues or lymph nodes.
† Amount of MAP was usually given in wet weight, but it was not always specified.
‡ Median time post infection when symptoms started, unless otherwise stated.
[$] Age when entering herd with infectious animals.
[d] "Natural": exposure was achieved by maintaining the animals in very heavily contaminated environments in which infected donor animals with clinical disease were continuously housed with the recipient animals.
Abbreviations: DPI= days post inoculation; IV = intravenous; Mo=months; NA=not assessed; yrs=years

Infection with MAP

Ingested MAP are transported through the intestinal wall via M-cells in the gut-associated lymphoid tissue (Momotani et al., 1988; Sigurðardóttir et al., 2004). The M-cells transport MAP to the subepithelial tissue where they are taken up by macrophages (Momotani et al., 1988). Degradation or division of MAP may then occur, and hence persistence and proliferation may occur if degradation does not take place (Sigurðardóttir et al., 2004). Activation of cell-mediated immune mechanisms is considered important for control of MAP proliferation. A Th1 population of CD4+ cells produce cytokines such as interferon-gamma (IFN-γ), and this contributes to activation of the macrophages and dominance of the cell-mediated immune response. By contrast, a Th2 subpopulation of CD4+ cells produce cytokines important in humoural immune reactions (Sigurðardóttir et al., 2004; Stabel, 2006). The mechanisms determining whether cell-mediated Th1 controlled processes are overtaken by humoural Th2 dominated mechanisms are still poorly characterised, and further research is required to identify triggers resulting in loss of control of an apparently latent MAP infection.

It is believed that there are several phases of infection. Early in the infection control is achieved by macrophages activated by IFN-γ and tumour necrosis factor-alpha (TNF-α). Clearance may or may not occur, because some inactivated, persistently infected macrophages remain. These macrophages are contained in granulomas, which are characteristic for MAP at pathological examinations. It is speculated that sporadic decay of macrophages within these granulomas accounts for transient bacterial shedding, although the animal is able to control the infection. The shedding might stimulate humoural immune response and production of IgG1 antibodies. This results in a progression of infection, because the cell-mediated immune reactions are no longer capable of controlling MAP proliferation. Occurrence of IgG1 antibodies indicates loss of control, as the protective immune reactions are suppressed and MAP spread in the animal. Extreme thickening of the intestines is caused by increased granuloma formation. A protein-losing enteropathy takes place, and the animal may suffer from diarrhoea, emaciation and eventual death (Coussens, 2001; Stabel, 2006). At present, it is suggested that as persistently infected macrophages increase in number in subepithelial intestinal granulomas, a low, but constantly increasing, stimulus of cytokines decreases pro-inflammatory and cytotoxic processes, leading to an increased number of MAP. It is also possible that persistently MAP-infected macrophages lack co-stimulatory activity, and that this leads to the development of regulatory T-cells, which leads in turn to the reduction of pro-inflammatory responses (Buza et al., 2004; de Almeida et al., 2008). MAP infections are primarily confined to the intestines and lymph nodes associated with the intestinal tract, but in animals with disseminated infection, MAP can also be found in livers, kidnes, lungs, heart and retropharyngeal, popliteal and prescapular lymphnodes (Antognologi et al., 2008).

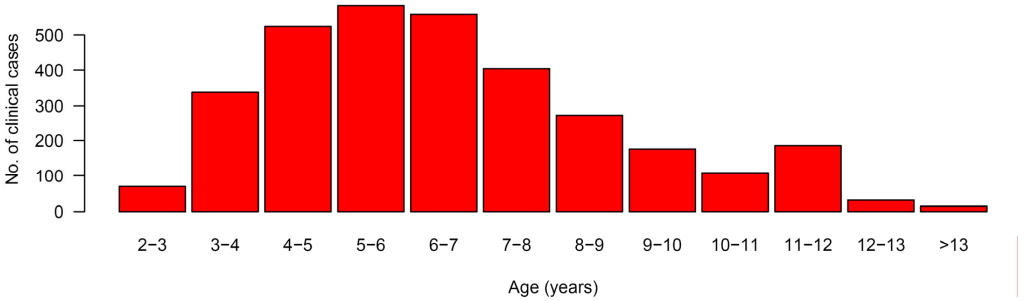


Fig. 1.2. Age distribution of clinical cases of paratuberculosis among 542 Victorian dairy herds participating in a control programme in Australia in 1992–2002. Modified from Jubb and Galvin (2004).

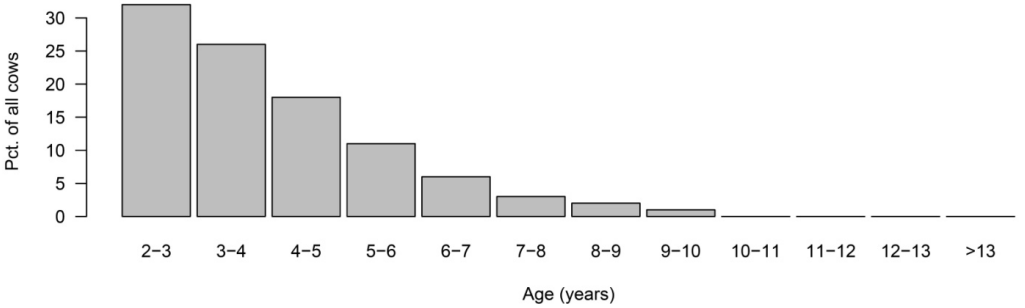


Fig. 1.3. Age distribution of all dairy cows in Denmark on December 31, 2005 (based on data from the Danish Cattle Database extracted Feb. 1, 2006).

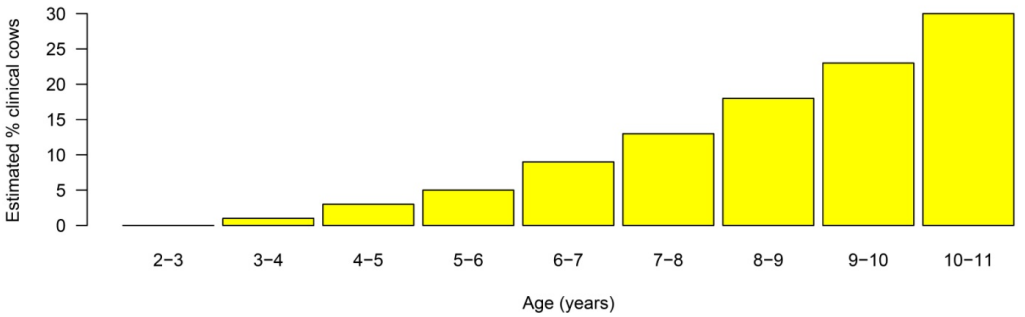


Fig. 1.4. Estimated expected age-distribution of cows with clinical disease among all cows recorded with clinical MAP infection, corrected from the total number of cows recorded with clinical MAP infection. The figure assumes that the age distributions of dairy cows in Denmark in 2005 and Victoria, Australia in 1992–2002 were the same. A clinical case only occurred once in each age-group.

Incubation period

The incubation period of MAP infections is not well characterised. This is primarily because there is no standardised infection model (Hines et al., 2007) and, at present, limited insight into mechanisms triggering immune shifts from pro-inflammatory to anti-inflammatory

responses (Coussens, 2004). However, clinical cases can occur at any time in life (Doyle and Spears, 1951; Jubb and Galvin, 2004). The age distribution of clinical cases in herds in a control programme operated in Victoria, Australia, is shown in Fig. 1.2.

The general age distribution in the Victorian cattle population was not given, but for comparison the age-distribution of all cows in Denmark on December 31, 2005 is shown in Fig. 1.3 Overlaying of the Victorian data with the Danish age-distribution suggests the distribution of cows with clinical disease shown in Fig. 1.4 This estimate is obviously an oversimplification, but it may be the best current estimate of the relative distribution of clinical MAP infections in different age groups. It appears that the proportion of cows suffering from clinical disease rises with increasing age. The figure assumes that the age distributions of dairy cows in Denmark in 2005 and Victoria, Australia in 1992–2002 were the same. If the assumption is sound, and if animals are most susceptible to infection as calves, it is safe to conclude that the incubation period can vary from a few months to the life-time of an animal.

1.4. Effects of MAP infection

Negative effects

Adverse effects of MAP infection have an impact at various levels: internationally, nationally, regionally and at animal- and herd-levels. Only herd- and animal-level effects will be discussed here (Fig. 1.5). Infected cows can, as a result of progressing infection, develop reduced milk production (Benedictus et al., 1987; Kudahl et al., 2004; Raizman et al., 2007a; Lombard et al., 2005; Hendrick et al., 2005), reduced body weight (Benedictus et al., 1987; Hutchinson, 1996), reduced fertility (Merkal et al., 1975; Johnson-Ifearulundu et al., 2000; Raizman et al., 2007a), and, ultimately, may be prematurely culled (Merkal et al., 1975). Infection can also be associated with other diseases, such as pneumonia (Raizman et al., 2007b) and mastitis (Baptista et al., 2008). However, the mechanisms behind these associations are not clear. Baptista et al. (2008) showed that in most cases increased somatic cell counts occurred prior to detection of antibodies to MAP in milk, but frequently the MAP antibodies occurred prior to increased somatic cell counts. There is therefore no clear causality, and the data suggest that impaired immune mechanisms in general may result in a higher risk of developing mastitis and in progression of MAP infection. It may be speculated that similar relations exist for other diseases, but this has yet to be explored. Be that as it may, the effects associated with MAP infections can be plentiful, although the presence of a MAP infection is not obvious.

Bacterial shedding affects the herd mainly because the increased bacterial load results in an increased risk of new infections and maintenance of the infection in the herd. Therefore, although the infectious cow is not necessarily affected herself, she may affect the herd by facilitating MAP spread. Control of MAP should therefore focus on these animals.

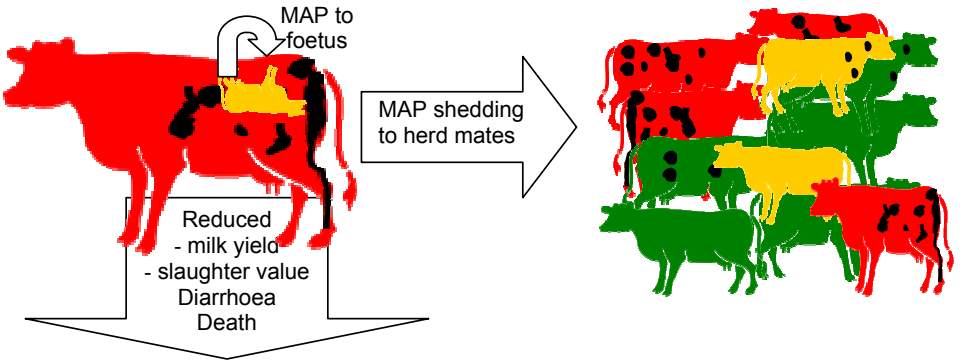


Fig. 1.5. MAP infections can have effects on multiple levels, because an infected cow can be affected by the infection leading to reduced milk yield, diarrhoea and death. She can also transmit MAP bacteria to susceptible animals resulting in their infection (shown by black spots representing faeces containing MAP). Furthermore, animals from a herd with infected individuals cannot be certified free of MAP infection, and the whole herd may have a reduced value for trade of live animals. The herd shown in this picture represents potentially non-infected animals (green), likely MAP-infected (yellow), likely MAP-infectious cows (red) and MAP-affected (red animal upside-down).

Lack of negative effects

MAP infection need not have an impact on body condition (McKenna et al., 2004), and cows which eventually become affected by MAP infection may actually produce more milk than their herd-mates in earlier lactations (Wilson et al., 1993).

1.5. Control of MAP infections

Several bovine control and eradication programmes have been run in Denmark. Targets have included *M. bovis*, the causative agent of bovine tuberculosis, bovine leukaemia virus, which causes enzootic bovine leukosis, and bovine herpes virus-1, the cause of infectious bovine rhinotracheitis. This section commences with a general discussion of prerequisites for the control and eradication of infectious diseases. This is followed by a brief description of the principles and practicalities of the Danish bovine tuberculosis programme and an overview of a selection of national paratuberculosis programmes. These parts of the thesis, in combination with information presented in Chapters 2–4, lead to an examination of the establishment of a Danish control programme on bovine paratuberculosis. The bovine tuberculosis programme was chosen because there are some similarities between MAP and *M. bovis* infections: a) detection of infected animals is hampered by imperfect tests; b) test-and-culling may be feasible if all reactors are culled, but a sensitive test is required, and some test-positive cows may never be affected by the infection; c) efficient breaking of transmission routes is essential if the infection is to be controlled effectively; d) a long-term

strategy is required to control and eradicate the infection, and this strategy will require the involvement of stakeholders at many levels, and particularly from industry. Only details regarding diagnosis, transmission, and control will be discussed here; no sociological issues will receive attention.

Infectious disease control and eradication

Infectious disease programmes can have different aims; they can also use different tools to reach these aims. The aims typically fall in three general categories: 1) eradication; 2) control; and 3) do nothing. Decisions about which aim to adopt are usually based on the effects caused by the infection, and on their impact on human or animal health, or the trade of live animals or animal products in specific regions. Therefore, decision-making can relate to different levels, depending on the potential impact of the infection. There is controversy over the definitions of "eradication" and "control" (Yekutiel, 1980). This controversy partly originates from the need to devise precise definitions that also have practical use. Therefore, Yekutiel (1980) has suggested the use of the Andrews-Langmuir definition (Andrews and Langmuir, 1963): "*Control is the purposeful reduction of specific disease prevalence to relatively low levels of occurrence, though transmission occurs frequently enough to prevent its permanent disappearance; eradication ditto but to the point of continued absence of transmission within a specified area*". Martin et al. (1987) are more specific, describing control as "*any effort directed toward reducing the frequency of existing disease to levels biologically and/or economically justifiable or otherwise of little consequence*". These definitions may still be rather abstract, but they are nevertheless useful. Yekutiel (1980) emphasises that although there are conceptual differences between the different terms, the key requirement is to be specific about the objectives of a programme. He also points to a number of preconditions (Table 1.2) which should be present to eradicate an infectious agent in practice. All of these preconditions are worth noting, but they may also differ from one infection to another, and some may be of minor importance in some programmes.

Table 1.2. Preconditions for eradication of an infectious agent in practice according to Yekutiel (1980).

P1: Existence of a main tool for effectively breaking transmission. It should be simple in application and relatively inexpensive.
P2: Case detection and surveillance in late stages of the programme should be possible.
P3: The infection should be of recognised socio-economic importance.
P4: Justification for eradication rather than control should be provided.
P5: Resources, including financing, administration, manpower and health services should be present.
P6: There should be no major socio-ecological factor which can negatively affect the programme.

Control of bovine tuberculosis in Denmark

Control and eradication of bovine tuberculosis was carried out in Denmark without accurate diagnostic tests. Financial resources were initially sparse and central coordination was not in place. At first only P1 seemed partly in place, although no major socio-ecological factors (P6) appeared to affect the programme. However, many infected cows tested positive without being affected by the *M. bovis* infection, so clearly the diagnostic test was not very specific for "cows affected by *M. bovis* infection". Therefore, the "Bang-method" (Bang, 1892; 1908; 1928) was attractive, because although repeated testing was required to detect infectious animals, there would then be no requirement to cull reacting animals as long as transmission of *M. bovis* was brought to an end. Briefly, the Bang-method included: a) cattle with udder tuberculosis or other evident signs of tuberculosis should be culled; b) whole-herd testing was carried out regularly using tuberculin skin testing; c) test-positive cattle should be separated from test-negative cattle; d) milk from the reactors should be pasteurised (at 80^{o}C) prior to use for feeding calves and swine; e) calves born from test-positive cows should be removed immediately from their dam and fed pasteurised milk or unpasteurised milk from test-negative cows. This approach might have significantly reduced the prevalence of reactors. However, because the control scheme was not systematically organised, control was successful only in some herds (Tønnesen, 1971).

Eradication of bovine tuberculosis from Denmark

It was only when the 1932 law on control of tuberculosis in cattle and swine from (Anon., 1932) was implemented that the number of "cleaned herds" approached 100% in other areas than the island Bornholm, which had no reactors from 1932 (Fig. 1.6).

This law introduced: 1) annual finances to spread information on the effects and transmission of *M. bovis* in cattle and swine; 2) financial support for diagnostic testing on a voluntary basis, but with a requirement to separate test-positive animals and test-negative animals if testing was subsidised; 3) recording of test-negative herds in official registers by Danske Mejeriers Fællesorganisation (Danish Dairy Board); 4) options for establishing trade restrictions on live animals from areas with high test-prevalences if more than 90% of an island's herds were free of infection (the "90%-rule"); 5) options for mandatory testing on an island if more than 90% of the remaining herds were free of infection; 6) subsidised culling of *M. bovis*-infected cattle with affected lungs (Anon., 1932). The implementation of this law, combined with pressure from the dairies to which the farms were delivering milk, appeared to be a necessary step in freeing all herds of *M. bovis*. It was demonstrated early on that, although the principles were efficient in individual herds, additional tools would be required to eradicate the infection from whole regions. For this purpose, the 90%-rule, also called Article 5 of the Law of June 26, 1932, proved efficient for control (and subsequent eradication) of *M. bovis* on a number of islands after the concept had been proved sound on the island

Bornholm. At the end of 1934, 11 islands with a total of 5,915 herds and 72,900 animals were practically free of reactors (Christiansen, 1935). The island certification thereby served several purposes: a) demonstration of proof of concept; b) stepwise eradication; and c) sources of replacement cattle with high likelihood of being free of *M. bovis* infection. The programme was in principle voluntary, except for culling of animals with open tuberculosis, but the dairies, most of which were farmer cooperatives, could decide on mandatory testing and culling in late-stages of the control phase to achieve eradication from their region. The process also showed that in areas that were not geographically isolated, such as Jutland, the process – , from the initiation of a systematic approach to a low prevalence of herds with test-positive animals – could take longer (Fig. 1.6).

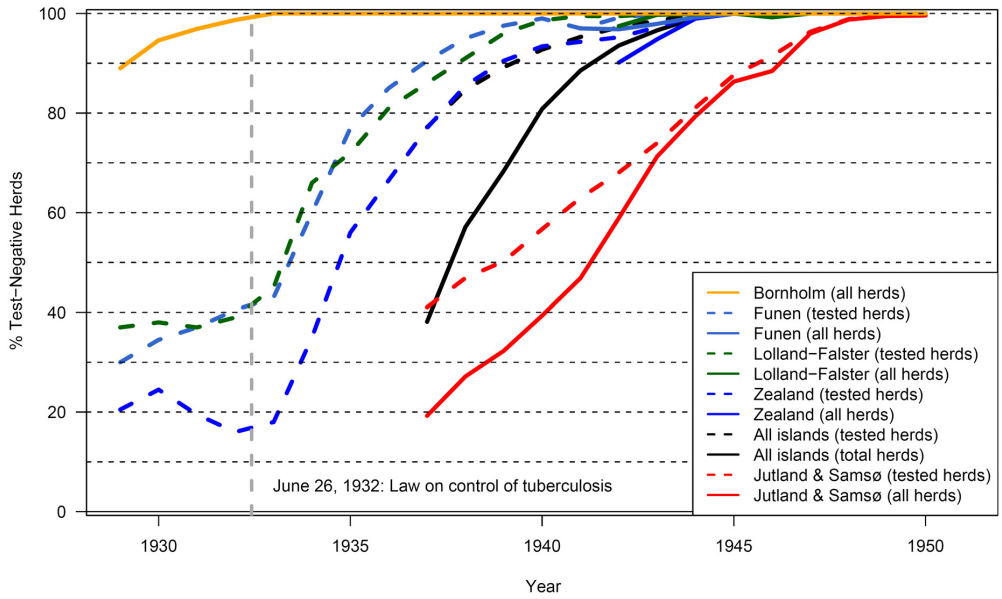


Fig. 1.6. Developments in prevalences of bovine tuberculin test-negative herds in Denmark 1929–1950. Assembled from data reported in Christiansen (1941) and annual reports of the Danish Veterinary Administration (Hansen, 1930; 1931; 1932; 1933; 1941; Nielsen, 1942; 1943; 1944; 1945; 1947; 1948; 1949a; 1949b; 1951; 1952). “All islands” includes Bornholm, Funen, Lolland, Falster, Zealand and other smaller islands, but not Samsø.

Bang (1928) observed that although a method referred to as the “American method”, where all reactors were culled, would be effective, it would also be very expensive and lead to loss of genetic potential. He also pointed out that the “Ostertag method” used in Germany, which relied on “carefully organised clinical examinations” with subsequent culling, was not sensitive enough to detect all animals with open tuberculosis. Therefore, the Bang method emerged as the most cost-effective method of control and eradication of *M. bovis* from cattle herds.

MAP control programme in the Netherlands

Several programmes have been established to control MAP infections in other countries. Benedictus et al. (2000) provide an overview of various approaches that have been attempted worldwide and programmes which have been carried out in the Netherlands. A programme initiated in 1998 had the following objectives: a) to reduce consumer's exposure to MAP; b) to minimise economic losses caused by MAP; and c) eradication of MAP from ruminant farms. Although considerable resources were put into the programme, only 473 of approximately 23,000 dairy herds were at "MAP-free" level by July 2005 (Weber et al., 2008a). A likely explanation is that annual whole-herd faecal culture of pooled or individual samples was required to obtain this status, and this test was not very sensitive or cost-effective and therefore not very appealing.

A new component was added in January 2006, aimed at reducing the concentration of MAP in bulk milk. By July 2008, approximately 15,000 herds (75% of Dutch dairy farmers) were included in the programme, with approximately 500 herds in the certification part, where the aim is low-risk trade of animals (Weber, 2008). The diagnostic test used in the milk quality assurance part is a commercial milk ELISA (the Pourquier milk ELISA, IDEXX, Montpellier, France) that is used for whole-herd testing. The cut-off in the test is higher than the one recommended by the producer to avoid false-positive reactions. Furthermore, herd-owners can choose to have ELISA-positive animals confirmed by faecal culture. A risk of the Dutch milk quality programme is that the sensitivity of the diagnostic test is reduced by increasing the cut-off and making confirmatory testing an option. Thereby, farmers with a few reactors may think their herds are free of infection, and, because test-negative herds are only tested every second year, the infection may spread in the meantime. Therefore, risk communication is of the utmost importance in reducing the risk of farmers, with a false feeling of security, believing that their herds are free of MAP infection due to lack of positive test results.

Control programme for MAP in the United States of America

In the US the Voluntary Johne's Disease Herd Status Program (VJHSP) was established in 1998 to set guidelines for identifying herds of low risk of MAP infection (Bulaga, 1998). The guidelines can be adopted by individual states as minimal requirements for participation in the programme, but more strict rules can also be applied. The objectives include minimal standards for certification of low-risk herds, and organised approaches to within-herd MAP control (Carter, 2007). Key elements of the programme are education of herd-managers, the establishment of management strategies to reduce the spread of MAP, and herd-classification based on diagnostic test results. Validation is secured through the National Johne's Disease Demonstration Project, which aims at evaluating long-term effectiveness and feasibility of proposed control strategies.

MAP control and certification programme in Australia

Australia has had to adapt their approach to suit a country in which there are huge regional differences in the incidence of MAP. MAP is rare in Western Australia but endemic in in the states of Victoria, New South Wales, Tasmania and South Australia, and different states had separate programmes until 1996. However, concern about the spread of MAP led to a national approach from 1996. The national strategic goals include: a) reduction of MAP contamination of farms and farm products; b) protection of non-infected herds and regions; and c) reduction of any social, economic and trade impacts of MAP. The programme includes both voluntary herd-certification schemes based on diagnostic testing as well as an option to have a risk score based on the management of calf-rearing. Control programmes in herds with endemic MAP infection include annual testing using ELISA of animals >4 years of age, subsidised culling of test-positive animals, and a calf-rearing plan aiming at reducing the transmission of MAP from adult animals. Education and communication are key factors, with a key message being “protect the next generation” (Kennedy, 2007).

MAP control and surveillance programme in Japan

Paratuberculosis has been a notifiable disease in Japan since 1971. All cattle detected with the disease are culled, and their farms are monitored until repeatedly test-negative. However, the prevalence of detected MAP cases has increased in the past 30 years (Kobayashi et al., 2007) irrespective of the fact that the disease has been notifiable, and there are no indications that the prevalence will decrease.

Plans for MAP certification in Finland

Finland may have a low level of MAP infections. It is therefore keen to certify herds free of MAP (Kulkas, 2007). However, the herd size is considered too small to obtain samples that would reliably certify herds as free of MAP infection (Kennedy and Nielsen, 2007).

Challenges in existing MAP control programmes

More programmes exist throughout the world, but the programmes above cover the main purposes specified by Yekutiel (1980): control and eradication. Many countries have not specified a set of strategic goals on MAP. The focus is generally on the choice of test and test strategy. These relate primarily to precondition P2, although P1 and P5 are generally considered more important in MAP programmes. Information and communication are parameters pivotal to P5. There is still debate as to whether the infection should be considered of “socio-economic importance” (P3). This debate is due in part to the long-term development of infection, where the negative effects are only seen after several years. Furthermore, difficulties in implementing measures to break transmission for all animals in a herd suggest that a risk-based approach, similar to the approach Bang suggested for bovine tuberculosis, should be used.

The chronic nature of MAP infections and the long-term development of adverse effects require that specific aims are identified for specific phases of control and eradication schemes, and that measures are established to monitor developments in different phases of the process. Control of MAP infections can theoretically be achieved without diagnostic tools by breaking the routes of infection. However, monitoring of the prevalence and detection of MAP-infected, infectious and affected animals can be used to guide and optimise the process. A characterisation of a) the performance of available diagnostic tools, b) their use in practice; c) prevalence estimation options; and d) risk factors will be given on the basis of currently available literature. The findings of the research underpinning this thesis are described in the following three chapters. Subsequently, the design of MAP control and surveillance schemes in Denmark is discussed.

2. DIAGNOSIS

2.1. Diagnostic tests

Diagnostic testing is usually carried out to inform subsequent decisions. These decisions may be about culling or therapy of the individual animal, or the establishment of a control plan for a herd of the kind based on a prevalence estimate. The results of testing can also be used to establish a prognosis. For chronic MAP infections, the prognostic properties are of particular importance, because while therapy is not an option for the individual, preventive measures can be established to avoid spread of MAP. However, a farmer may wish to ensure that an animal produces as much milk as possible prior to culling, and hence proper timing of the culling can become essential from a cost-effective point of view. Therefore, knowledge of the performance of a diagnostic test is essential from several perspectives. Test validations should therefore be carried out in relation to a specific purpose, to address the "fitness for purpose" criterion as endorsed by OIE (Anon., 2003).

Animal-level testing

Ante-mortem diagnostic tests used for paratuberculosis can be divided into agent-detecting tests and tests detecting cell-mediated or humoural immune responses (overview provided in Nielsen et al., 2001). Furthermore, gross and histopathological examinations can be performed for diagnosis of paratuberculosis, primarily post-mortem. The range of diagnostic tools used for diagnosis at animal-level has been wide, but since 1990 reports from test-validation studies have primarily included indirect ELISA for the detection of antibodies in milk and serum, ELISA for the detection of IFN-γ in whole-blood, or the detection of MAP in faecal samples using bacteriological culturing (test accuracies reviewed in Nielsen and Toft, 2008 I). Prior to 1990 the complement fixation and agarose gel immunodiffusion tests were commonly used, but they have been reported to give lower accuracy than indirect ELISA (Sockett et al., 1992). Detection of MAP in milk samples by culture (Giese and Ahrens, 2000) or IS*900*, or other specific genetic elements in samples of milk (Giese and Ahrens, 2000), faeces (Eamens et al. 2000) or blood (Buergelt and Williams, 2004) is also possible, but the accuracy of these methods has either not been reported or been reported to be inferior to those of conventional culture methods (Eamens et al., 2007). Histopathological evaluation of tissue is also an option (Buergelt et al., 1978; Pérez et al., 2005), but is rarely carried out in routine diagnostics for cattle. Whitlock et al. (1996) obtained 5–46 tissue samples from each of 171 cull cattle with a previous positive faecal culture and found a correlation between the number of sampled tissues and the probability of a positive diagnosis. They suggested that up to 100 sites per animal should be sampled to detect MAP in cows without clinical disease.

Herd-level tests

Herd-level tests can be used to classify herds as "infected" or "non-infected" or "low-risk" and "high-risk". These classifications can be used to reduce the risk of transmission between

herds. A herd-level test result can also be used as an indicator of whether a control scheme should be established, or to assess prevalences in infected herds.

Herd-level tests include individual and pooled tests aggregated to herd-level, as well as culture of environmental samples. Test-strategies using these tests are typically developed to estimate whether a herd is free of MAP infection (Bulaga, 1998; Kalis et al., 2000). However, because of the imperfect sensitivity of individual tests (Nielsen and Toft, 2008 I), and because pooling erodes sensitivity (Kalis et al., 2000; Wells et al., 2002), such test-strategies cannot be used for the purpose of certifying herds as free of infection. They may be used to confirm the infection in a farm. No studies have provided unbiased estimates of the accuracy of herd-level tests, although several attempts to do this have been made (e.g. Kalis et al. 2000; Wells et al., 2002; 2003). Therefore, unless historical information is also included, herd-level tests can only be used to classify the probability of a herd having a high or low prevalence of MAP-infected animals (see Chapter 3; Martin, 2008).

Usually, herd-status has been defined on the basis of the culture of faecal samples from individual adult animals (Whitlock et al., 1992; Kalis et al., 2000; Wells et al., 2002; Lombard et al., 2006). Lombard et al. (2006) demonstrated that 38 (76%) of 50 herds classified as infected by faecal culture and 61 (76%) of 80 herds classified as infected by ELISA were also positive using a culture of approximately 5 environmental samples per farm. Raizman et al. (2004) detected 78% of herds known to be infected through a control programme using environmental culture. Low-prevalence herds (Wells et al., 2003), or herds where young-stock are infected without MAP shedding among the adult cows, are at high risk of not being detected using such herd tests, however. Detection of MAP by culture (Grant et al., 2002), MAP genetic elements using PCR (Grant et al., 2002; Stephan et al., 2002; Bosshard et al., 2006) and antibodies using ELISA (Nielsen et al., 2000, Böttcher and Gangl, 2005; van Weering et al., 2007) have also been used on bulk tank milk samples. A low correlation between the within-herd prevalence based on individual milk ELISA and antibody levels in bulk tank milk ELISA have been found in several studies (Dussol, 2004; Böttcher and Gangl, 2005). Herthnek et al. (2008) reported a low correlation between the detection of MAP DNA in bulk tank milk and MAP detection in environmental samples, suggesting that while bulk tank milk PCR is not sensitive for detecting infected herds, it may be able to detect MAP DNA in milk, if such DNA is present.

2.2. Purpose of testing

The chronic nature of MAP infections and their multiple potential effects require the decision maker to decide on both the purpose of testing and the subsequent use of the test results. The objectives of testing might be: a) estimation of prevalence to determine whether a control programme should be established either to study risk factors or to perform simulation studies in which prevalence estimates are required; b) to predict shedding from individuals; c) to predict time-to-occurrence of drop in milk production, or d) certification of herds or

animals as free of MAP infection. Three conditions 1) MAP-infected, 2) MAP-infectious, and 3) MAP-affected, were identified as useful target conditions in connection with the above-mentioned objectives (Nielsen and Toft, 2008 I). For example, MAP-infected may primarily be of relevance for prevalence estimation, MAP-infectious when transmission should be avoided in a control scheme, and MAP-affected when individuals are affected to a degree that their performance is reduced.

These target conditions can assist with efforts to establish the purpose of the intervention and the reason for testing. MAP-infected animals carry the bacteria intracellularly, but some infected animals control the infection through pro-inflammatory immune reactions, whereas others proceed to MAP-infectious or MAP-affected stages. MAP-infectious animals shed sufficient bacteria to result in an infection in a susceptible animal, whereas MAP-affected animals can have chronic or intermittent diarrhoea, weight loss, reduced milk yield, reduced fertility or other effects, demonstrating that the animal is actually affected by the MAP infection even though this would not necessarily be obvious to a clinician. Weight loss, reduced milk yield, and the like, are short-term adverse effects. Infectiousness is a long-term adverse effect: it has a negative effect, not necessarily on the infectious animal, but on its herd-mates.

Further target conditions may be specified for special circumstances. For example, “Occurrence of MAP in milk” might be useful to some decision makers, such as those in the Netherlands involved in a milk quality assurance programme aiming to reduce the risk of high level of MAP in milk delivered for consumption (Weber and van Schaik, 2007; Weber et al., 2008a).

2.3. Diagnostic accuracy and costs of testing

ELISA versus agent-detecting tests

It is frequently stated that agent-detecting tests are “better” because antibodies are only produced late in infection (Collins, 1996; Stabel, 1997; Clark et al., 2008). However, it should be considered what they are “better at”. For example, Clark et al. (2008) found that an IS*900* direct PCR test was better at detecting MAP in a sample that was defined as positive when MAP had been cultured from the same sample than an ELISA test was in detecting the cow from which the sample was obtained as a MAP shedder. In this case, the study objects in the two samples were different: the PCR was carried out to detect MAP in the faecal sample, whereas the ELISA was carried out to detect a cow shedding MAP. Furthermore, in half of the study population, culling based on ELISA was carried out. Hence, cows most likely to have had immune responses were removed before the study. The design of this study would result in both selection bias and misclassification bias, neither of them favouring the ELISA test. It would therefore be inappropriate to conclude that the agent-detecting test was better than the immunity-based test. Studies comparing accuracies between ELISA and agent-detecting tests remain sparse.

A review of reported accuracies of diagnostic tests did not support the statement that agent-detecting tests are better than ELISA at detecting MAP-infected animals (Nielsen and Toft, 2008 I). The lack of a reference method for evaluating and comparing diagnostic tests is a major obstacle, and the majority of studies have therefore defined faecal culture methods as the reference method in the study design, thereby excluding the possibility of an assessment of the accuracy of faecal culture. However, opportunities for sampling and cost are the primary factors affecting choice of test, and if faecal culture and ELISA have accuracies in the same range (Nielsen and Toft, 2008 I), then milk ELISA is an obvious choice. In 2006 median cost/test in the USA was reported to 17 USD, 5 USD and 6 USD for faecal culture, serum ELISA and milk ELISA, respectively, excluding sampling costs. In Denmark, the costs for faecal culture can be estimated to 27 EUR / faecal sample (incl. sampling costs) and 3 EUR / milk sample (incl. sampling costs) if 100 animals are sampled, using the unit costs provided in Sergeant et al. (2008 VIII). These prices, in combination with ease of sampling, would favour use of milk ELISA as long as its accuracy is high enough for control purposes. Ease of sampling appears to be particularly important if frequent testing is carried out, as is explained in the following sections.

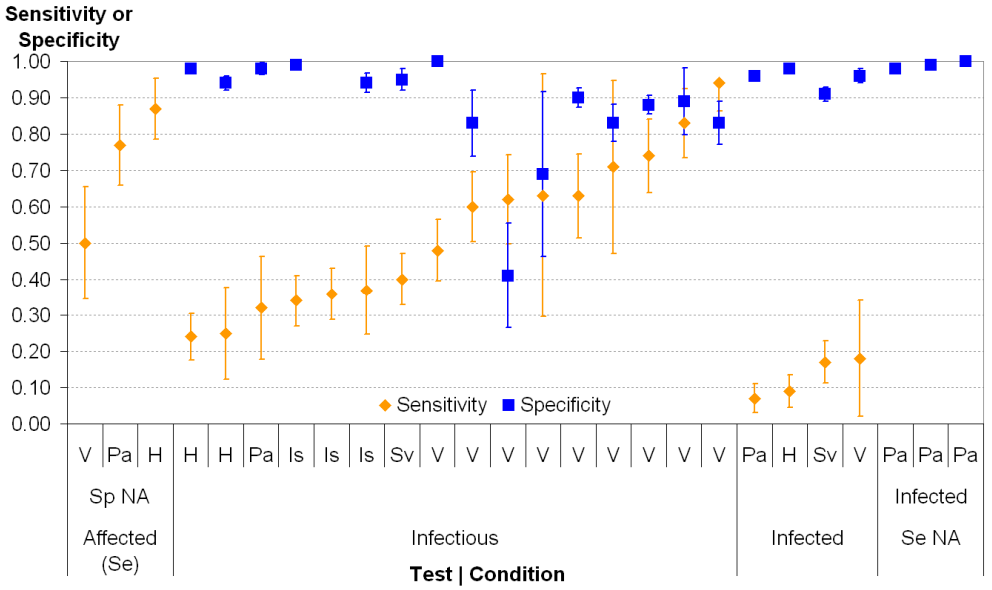


Fig. 2.1. Reported sensitivities and specificities of serum ELISA used for affected, infectious and infected cattle, respectively (based on data summarised in Nielsen and Toft, 2008 I). Tests: Pa=Parachek, H=Herdchek; Is=IDEXX Sweden; V=Various. Se NA=sensitivity not assessed; Sp NA=specificity not assessed; Se NA = Sensitivity not assessed.

Detection of MAP-infected and MAP-infectious animals

The reported accuracies of ELISAs are highly variable, and usually high sensitivity has resulted in low specificity and vice versa (Fig. 2.1; Nielsen and Toft, 2008 I). Therefore, it is

important to know the estimates of the specific ELISA used, and to use it according to these estimates. Because MAP infections are chronic, the sensitivity and specificity with which MAP-infected animals can be detected increases with age (Nielsen and Toft, 2006 IV). The sensitivity of the in-house ELISA used in Denmark from 1999 to 2008 (Nielsen, 2002) appeared to increase from approximately 0.06 at 2 years of age to 0.50 at 5 years of age, whereas the specificity decreased from 0.997 to 0.93 across the same age-span (Nielsen and Toft, 2006 IV). The detection of MAP-infectious animals appears to be affected more by test frequency than it is by age (Nielsen and Toft, 2006 IV). Most animals shedding MAP will develop antibodies (Nielsen and Ersbøll, 2006 II), and therefore the period from development of antibodies to detectable shedding is of interest in connection with the detection of MAP-infectious animals. Among animals classified as high shedders, i.e. cows with repeated positive faecal culture and where a minimum of one sample has ≥50 colony forming units, 70% were ELISA positive on the date bacterial shedding started (Nielsen, 2008 V). Only 5% of transient shedders, i.e. cows with a positive faecal culture followed by a minimum of 4 negative faecal cultures, were ELISA-positive on the date of first detected bacterial shedding. If the high shedder had been tested one year prior to start of bacterial shedding, only 27% of the animals would have been ELISA-positive, whereas approximately 60% would have been detected using quarterly samples. The data thus suggested that frequent testing is required if MAP-infected animals are to be detected before they become MAP-infectious.

Specificity considerations with repeated testing

Frequent testing will result in a higher cumulative frequency of test-positives, some of which will be false-positives. It is important to consider, what a “false-positive” is in this context. Three to four years before they become high shedders, approximately 20% of the animals were detected as ELISA-positive (Nielsen, 2008 V). The pattern can be reproduced with a different commercial ELISA (Pourquier ELISA, IDEXX, Montpellier, France). The commercial ELISA appeared to have both a higher sensitivity and higher specificity (Fig. 2.2). The data suggest that antibodies can begin to develop many years before onset of detectable shedding. This is in keeping with the conceptual development of MAP infections, where anti-inflammatory immune reactions begin an early battle with the pro-inflammatory responses and slowly come to predominate (Coussens, 2001). However, had the animal that would later be a high shedder been test-positive 3 years before being detected as a high shedder, the probability of confirming the diagnosis using an agent-detecting test would have been very low (Nielsen, 2008 V), and she could be considered a false-positive in the sense that she was non-MAP-infectious. However, she would most likely be MAP-infected.

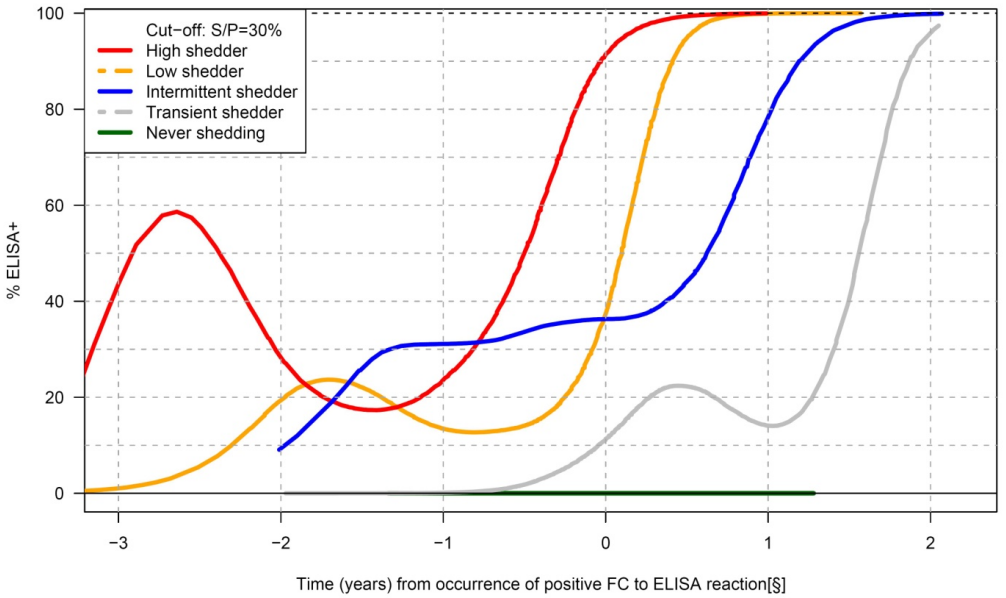


Fig. 2.2. Probability of testing positive in Pourquier ELISA relative to start of High, Low, Intermittent and Transient shedding of MAP. [§]For cows not shedding MAP, the median in a test-period was used as Time = 0 (Nielsen, unpublished data).

ELISA results on an ordinal scale

ELISA results are usually dichotomised, so as to categorise animals as test-positive or test-negative, for example, and to estimate sensitivity and specificity relative to a given target condition. However, they are measured on an ordinal scale, and hence dichotomisation results in loss of information. Therefore, decision-making can be enhanced if results are used on an ordinal scale. Collins (2002) demonstrated that the likelihood ratio for a positive result, i.e. the odds of disease at a given level of the diagnostic test, increase with increasing value of the ELISA. Toft et al. (2005 III) further divided animals into no MAP shedding, intermittent MAP shedding and continuous MAP shedding, added covariates for parity and stage of lactation, and demonstrated that the probabilities of bacterial shedding in the two shedding groups increased with increasing corrected optical density readings from the ELISA test. The results here also suggested that an ELISA value has to be interpreted as a probability of bacterial shedding rather than a dichotomised shedding, yes or no. This probability is often significantly different from 1, but the uncertainty associated with it can be reduced if covariates are incorporated and interpretation of multiple outcomes is permitted. Ultimately, the challenge is application of probabilities and communication of the results in practice.

Detection of animals affected by MAP infection

MAP-affected animals include animals where the infection affects the animal by, for example, reducing milk yield, reducing body weight and causing diarrhoea. Serum antibody ELISA has

been reported to have a high sensitivity in animals with clinical MAP infection (reviewed in Nielsen and Toft, 2008 I), and as a confirmatory test the Pourquier ELISA has been reported to have a sensitivity of 0.83 to 0.92 and a specificity of 0.998 (Weber et al., 2008b). The confirmatory diagnosis can be important, in connection with individual cows, in determining whether the correct clinical diagnosis has been made. However, it can also be important to determine the value of the ELISA as a prognostic test for MAP-affected animals. It can thereby be determined whether the infection has progressed to a stage at which culling should be advised. It was demonstrated that cows with fluctuating ELISA responses produced ~0.5 kg more energy-corrected milk than their ELISA-negative herd-mates (Nielsen et al., 2009 VI). They therefore appear not to be affected negatively by MAP infection and thus should not be considered as members of the MAP-affected group. Milk yields in cows with repeated positive ELISA-results were also higher than their test-negative herd-mates until approximately 100 days prior to the first positive test. Thereafter yield decreased significantly. Interestingly, cows with only one positive ELISA on the last test-date which were not confirmed positive by a second ELISA or were deemed fluctuating because that ELISA later turned negative, started to display reductions in milk yield 300 days before testing positive. These results suggest that in some cows, the infection progresses so quickly that the ELISA will only be positive after production loss has occurred. Farmers often cull these animals because of the production loss without waiting for the confirmatory test result. It should be noted, however, that the Danish recommendations are that at least two consecutively positive tests are required before a cow is deemed to be a candidate for culling due to MAP infection alone. Therefore, the results indicate that a combination of repeated ELISA and a reduction in milk production greater than expected can be a practical tool for decisions on culling. These results are concordant with the findings in Wang et al. (2006) where it was demonstrated that the accuracy of milk ELISA increased with decreasing milk yield.

To conclude, repeated milk antibody ELISA is needed to determine whether a cow is affected by MAP infection or merely infected with MAP. False-positive reactions can occur in cows that have tested positive previously.

Cross-reacting antibodies

Most ELISAs are not 100% specific, as cross-reacting antibodies to non-MAP *M. avium* infections may occur. Absorption using *M. phlei* should remove cross-reacting antibodies to non-*M. avium* infections (Yokomizo et al., 1985). However, infections with MAA, MAH and MAS may result in antibodies that cross-react with antibodies reacting to MAP. *M. avium* Serotype 2, an MAA strain, was frequently isolated from cattle in the past in Denmark (Jørgensen, 1978), but has not been isolated frequently recently. Whether this is due to changes in management or lack of testing is unknown.

False-positive test results

False-positive reactions resulting from cross-reacting antibodies to non-MAP infections are of particular concern in certification schemes. Therefore, confirmatory testing using an agent-detecting test is common in countries such as the US (Bulaga, 1998) and the Netherlands (Weber, 2008). However, not all false-positive reactions result from cross-reacting antibodies. Some reactors are MAP-infected, but not yet MAP-infectious. For example, only approximately 10–30% of antibody-positive cows were confirmed culture positive in Nielsen (2008 V). Because 1–7% of the animals that are ELISA positive would be considered non-infected ($Sp_{ELISA,\ MAP\text{-}infected}$ ~93–99%, Nielsen and Toft, 2006 IV), the majority of reactors would be expected to be MAP-infected. Because only 10–30% of the antibody-positive cows shed MAP, the majority would actually be in the phase prior to being MAP-infectious. These animals will go undetected and hence may be regarded by farmers as "free" of MAP-infection. Of course, they may become MAP-infectious and spread MAP in the herd for years where measures to reduce transmission are not established. Therefore, it is important to consider carefully the definition of "false-positive" in a given test scheme: is it meant to cover cross-reacting antibodies only, or does it also include MAP-infected animals which have MAP specific antibodies but are not yet MAP-infectious?

Summary of test interpretation

To summarise, the interpretation of ELISA results become more precise if a purpose of testing is specified and a target definition to be used in relation to this purpose is given. Milk ELISA has a low sensitivity for the detection of MAP-infected animals, but the sensitivity increases with age. In view of the generally unpredictable incubation period, timely detection of MAP-infectious animals requires repeated, continuous testing from 2 years of age and onwards. Results from continued monitoring can be used to predict bacterial shedding and milk production losses, but ELISA may fail to detect some infectious animals and some cows experiencing production losses. ELISA can be used effectively to confirm clinical disease, particularly because the results can be available within 1–2 weeks.

Milk antibody ELISA has the advantage of being able to detect many MAP-infected animals before they become infectious if testing is undertaken frequently. Furthermore, it is a relatively inexpensive test. Lastly, milk samples are easily obtained, which is a requirement if frequent testing is to be carried out.

Challenges in the use of diagnostics remain, because great uncertainty is associated with the tests, but for certain purposes the probability of detecting the target condition is sufficiently good.

2.4. Application of results in practice

Most of the animal-level tests require individual animals to be caught and sampled, requiring a workload that many farmers are not willing to accept on a regular basis, if frequent testing

is required. Furthermore, the handling of animals may cause stress to the animals unnecessarily, potentially resulting in adverse effects such as reduced milk yield. The use of routinely collected milk samples can reduce both the workload of herd managers and any undesirable impact on the animals. Somatic cell counts are already routinely made through various milk recording schemes, and systems to do similar routine diagnostic surveillance could be applied for diagnoses related to MAP.

Repeated test results plotted on an ordinal scale for a given cow over time, when presented graphically, can provide a decision maker with an overview of that particular animal. However, in a herd of 100 cows 100 antibody profiles could result, and these may not be easily categorised for use in a management system, since limited existing understanding of such results can be expected on many farms. Some herd-managers may wish to be given detailed information about the individual animal. Herd-managers are usually decision makers in the herd setting, but to make decisions they need to specify what desiderata decisions should be guided by. The aim of the Danish control programme on paratuberculosis is to provide tools for herd-managers to reduce the prevalence of MAP infections (Nielsen, 2007). Test-and-cull strategies are ineffective without changes in management to reduce transmission (Kudahl et al., 2007). However, risk-based control strategies with four annual herd tests using a milk ELISA can result in cost-effective reduction of MAP infections. For success here animals with a high risk of being infectious (High-Risk animals) need to be identified and transmission from these animals to susceptible calves needs to be avoided (Kudahl et al., 2008 XIII). The most infectious cows should be culled prior to next calving. There are thus three groups of animal to be identified: a) Low-Risk cows; b) High-Risk cows that can remain in the herd, but from which transmission should be avoided; and c) High-Risk cows that should be culled. For ease of communication, these animals have been termed “Green”, “Yellow” and “Red”, respectively. At high prevalences, the number of High-Risk cows deemed apt for culling may exceed the number of replacement animals available. It is important to note, then, that evaluation of antibody profiles can be useful for determining which animals are most likely to excrete high levels of bacteria and, hence, should receive highest priority in any culling. Additional groups of animals can be identified. Calves born of cows that are, or will become, High-Risk animals around calving are one such group. These calves should be evaluated on the basis of their dam’s profile, because of the risk of transmission *in utero*.

To summarise, then, frequent testing enables many MAP-infected animals to be detected before they become infectious. MAP-infected but potentially non-infectious cows pose a special challenge in daily management, but risk-based management with multiple risk-groups can be used.

3. PREVALENCE OF MAP INFECTIONS

Prevalence estimates for a particular infection can be used by decision makers to determine whether an intervention should be initiated, or to monitor the level of intervention over time. They can also be employed by scientists to study risk factors, run simulation studies, perform risk assessments, and so on. There are many reports on the prevalence of MAP infections (reviewed in Nielsen and Toft, 2009 VII). However, comparable and reliable prevalence estimates are generally not available in Europe, and many countries have not reported recent estimates. Comparable prevalence estimates may come to be required if international trade restrictions are put in place, or if joint international actions to control MAP are agreed. The main reasons for the lack of comparable prevalence estimates are variable definitions of the condition “paratuberculosis”, a lack of accurate tests to detect the condition, and poor study designs. Variable definitions are a natural consequence of the chronic nature of an infection with an incubation period usually lasting years.

Prevalence of MAP infections in Europe and North America

Going by European data, the best guesstimate of the between-herd level of MAP infection was >50%. Animal-level prevalences were guesstimated to approximately 20% (Nielsen and Toft, 2009 VII). Some countries may have a lower prevalence, but these estimates have not been reported in a form permitting them to be satisfactorily validated. Generally, prevalence estimates from non-European countries are also non-comparable and non-interpretable, although the designs of several studies have individually appeared reasonably sound. The apparent prevalence in the Maritime provinces of Canada based on serum ELISA was estimated to 2.6% among dairy cows (VanLeeuwen et al., 2001), but a prevalence corrected for sensitivity and specificity was not calculated. A study conducted five years later in the same area, based on culturing of tissues samples from the ileum and from mesenteric lymph nodes, revealed an apparent prevalence of 16.2% among abattoir cattle (McKenna et al., 2004). This discrepancy in prevalence estimates suggested that the ELISA had a very low sensitivity, and that the true prevalence could not be determined. Collins et al. (1994) recorded an apparent prevalence of 7.3% based on serum ELISA and tracked sensitivity and specificity to 0.509 and 0.949, respectively. However, the sensitivity estimate for ELISA was based on a target condition faecal culture positive (~MAP-infectious), whereas the specificity estimate was based on a target condition free of MAP infection. Thus two different target conditions were used, and a true prevalence of MAP-infected animals could not be estimated.

The lack of uniformly defined prevalence estimates hinders the comparison of prevalences across regions and countries. However, within-region or within-country estimates may be useful if the same test system is used in follow-up surveys. Failures to report prevalence estimates may occur as a result of political pressure from decision makers

in some countries, or in response to scant interest from scientific journals, or because too few researchers are willing or able to carry out and report the studies.

Prevalence of MAP infections in Denmark

Prevalence estimates from Denmark are few, and those that exist suffer from some of the flaws mentioned above: the same tests and target conditions have not been used; study designs have not been suitable for providing unbiased estimates of prevalence; and there has been a lack of reporting due to low priority. Bovine chronic diarrhoea believed to be associated with MAP infections was first reported from the island of Lolland in late nineteenth century or early twentieth century (Bang, 1906; 1909).

Multiple estimates of prevalence have been reported during the twentieth century, but most have been based on laboratory submissions (e.g. Jørgensen, 1972) or opinions (e.g. Andersen, 1965), and there are few objective, systematic investigations among the reports. The first systematic epidemiological study was conducted by Jørgensen (1965), who cultured MAP from mesenteric lymph nodes from 2 (0.45%) of 448 abattoir cattle <2 years of age, and 15 (2.3%) of 662 abattoir cattle >2 years of age. Bulk tank milk used for a survey of 900 Danish dairy herds revealed an average true prevalence of 47% in a stratified sample (Nielsen et al., 2000). However, that survey was impaired by use of an ELISA method that was not yet fully optimised. Selection of a cut-off with an ELISA value of 0.03 values higher than the one selected would have resulted in an overall estimated true prevalence of 13%. ELISA scores in the test used are usually in the range 0–2, and a small change can therefore have major effect on prevalence. Studies using an optimised method have never been reported in the scientific literature, but one was mentioned in a farmers' magazine (Nielsen, 2000). The study in question was carried out in 1999 using a bulk tank milk antibody ELISA (described in Dussol, 2004) with estimated herd-level sensitivity and specificity of 0.94 and 0.86, respectively. True prevalences at herd-level (Nielsen et al., 2004) were estimated to 85%, 86% and 80% among 2,709 herds located in Western Jutland, South Jutland and Southern Jutland, respectively. Furthermore, in another unpublished study examining 59 organic herds delivering milk to the same organic dairy in Jutland, 39 (66%) were found positive in culture of 6–18 environmental samples in 2006. If the sensitivity is assumed to be 0.78 (Raizman et al., 2004) and the herd-level specificity 1.0 (confirmatory IS*900* PCR was carried out on culture positive isolates), estimated true prevalence comes out at 0.85, which is in the same range as the bulk tank milk estimates for 1999 given above. These studies confirmed that the prevalence of MAP infection in Denmark is high, and indeed these estimates partly drove the decision to establish a control programme on MAP in Denmark (see Chapter 5).

Relative estimates of MAP prevalences

The derivation of absolute cow-level prevalence estimates is challenging (Nielsen and Toft, 2009 VII). However, as long as the same diagnostic method is used, relative estimates can be useful both for risk factor studies and for comparisons with other herds. Dichotomisation of ELISA values, which are recorded on an ordinal scale, can result in loss of information. A possible solution is to use the actual ELISA values to estimate the within-herd prevalence correcting the estimate for laboratory and cow-level effects using a mixture-model. Thereby, the problem of choosing a cut-off for the ELISA can be circumvented, and the estimates obtained will be more precise than they would be using a cut-off based method (Nielsen et al., 2007b IX). This method can be used to rank herds against each other, but point-estimates are less reliable due to the poor test-performance. The ranking can then be used for studies of risk-factors, irrespective of the absolute size of the estimate. Since diagnostic tests are in general poorly suited for prevalence estimation, it is necessary to use as much information as there is available to use these estimates in such investigations as risk factor studies and studies of the probability of freedom from infection.

Herd-classification based on imperfect tests

Ranked relative prevalence estimates can be obtained using a mixture-model; alternatively herds can be classified as low-risk and high-risk using herd-level tests (described in Chapter 2). However, test-negative may by be equated with “free of MAP infection” by some farmers irrespective of the fact that, due to imperfections in the tests, a test-prevalence of 0% does not necessarily suggest that the true prevalence is 0%. The test-prevalence in herds with the same true prevalence will be lower for herds with a low median animal age than it is for herds with high median age, because the probability of detecting MAP-infected animals increases with increasing age. Therefore, age-stratified sensitivity and specificity estimates can be used to estimate an age-corrected true-prevalence using the Rogan-Gladen estimator (Sergeant et al., 2008 VIII). This estimate can be transformed into a probability that a herd has a prevalence lower than a given design prevalence, or the population prevalence can be used as a design prevalence. In this way herds can be classified on the basis of their probability of having a prevalence that is lower than a pre-defined estimate. Five different test-strategies were evaluated: 1) milk ELISA testing of all lactating animals; 2) milk ELISA testing of young lactating animals (≤4 years of age); 3) milk ELISA testing of old lactating animals (> 4 years of age); 4) faecal culture of all lactating animals; and 5) serial testing with milk ELISA of all lactating animals and confirmatory testing using faecal culture of ELISA-positive animals. Whole-herd ELISA was identified as the most cost-effective method of detecting herds with MAP-infected animals and providing high probabilities (>0.9) that non-infected herds had a low prevalence. However, herds with <80 cows and herds with a young age-structure had too few cows to classify them with reasonable certainty (Sergeant et al., 2008 VIII). The results suggested that whole-herd ELISA can be used as a classification

strategy, but herds cannot be classified as free of MAP infection on the basis of single herd tests. Inclusion of historical test-information of the kind described by Martin (2008) could be an option if certification should become part of a programme in Denmark.

4. RISK FACTORS FOR TRANSMISSION OF MAP

Most MAP shedding occurs from cows, with MAP bacilli being ingested primarily via faeces. Calves are most susceptible to this kind of ingestion, and therefore the primary risk points are: a) the calving area, where new born calves are in close contact with their dams and potentially other cows; b) the use of milk and colostrum for milk feeding; c) calving of cows in late stages of MAP infection. Other factors resulting in calves having high levels of exposure to the manure of adult animals are also of great concern. Lack of solid evidence on when infections occur, the following long incubation period and the lack of an accurate diagnosis, makes studies of risk factors challenging. There have been no systematic reviews of risk factors associated with paratuberculosis, and the studies carried out have focused primarily on associations between risk factors and a given diagnostic test result rather than actually elucidating risks of MAP transmission.

Types of risk

"Risk" can be roughly divided into risk of exposure, risk of infection given exposure and risk that the infection affects the animal exposed to MAP (Fig. 4.2). Risk of exposure is typically related to management factors. If animals are not exposed to MAP, infections will not result, and the challenge is to keep MAP away from the herd. Risk of infection given exposure may be primarily genetically determined, whereas risk of being affected may be determined by both non-genetic factors (e.g. exposure dose, stress-related factors and age) and genetic factors (e.g. the ability to mount an effective MAP immune-response).

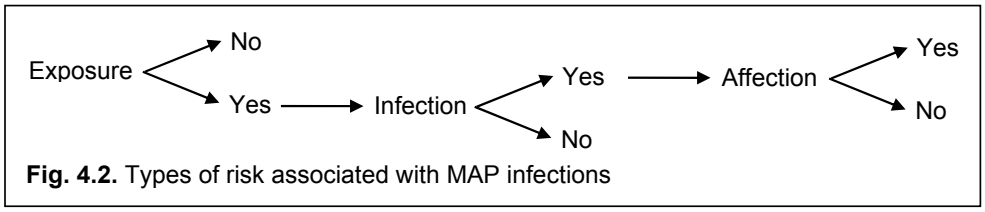


Fig. 4.2. Types of risk associated with MAP infections

Because exposure, infection and affection are not easily measured with current diagnostics, and because genetic and non-genetic factors may be correlated, the majority of risk-factor studies are difficult to interpret. Furthermore, many risk factor studies have been based on herd-level classification, with one or two test-positive individuals being used to define the herd-status. The meaning of the outcome variable in these studies is difficult to interpret in relation to the above-mentioned outcomes.

Host-related risk factors

Reported host-factors include breed for development of clinical disease (Cetinkaya et al., 1997), genetic variation in the ability to raise MAP antibodies (h^2=0.10; Mortensen et al., 2004; Gonda et al., 2006) and the combined ability both to raise antibodies to MAP and shed

MAP (h^2=0.10; Gonda et al., 2006). Age can be related both to risk of infection given exposure and risk of becoming infected. The risk of becoming MAP-infected given exposure is considered highest among young animals (Taylor, 1953). Increasing age has been associated with increased probability of detecting MAP infection (Nielsen and Toft, 2006 IV) because the infection is chronic. The risk of becoming MAP-infectious was highest between 2.5 and 5.5 years of age (Nielsen and Ersbøll, 2006 II). Older cows apparently have a higher probability of becoming affected by MAP infection than young animals (see Chapter 1), but the effect on milk yield may be similar irrespective of parity (Nielsen et al., 2006).

Environmental and management-related risk factors

Environmental and management-related factors are not easily separated, because specific management factors may result in an increased presence of MAP in the farm environment. Numerous studies have been carried out (Collins et al., 1994; Obasanjo et al., 1997; Cetinkaya et al., 1997; Johnson-Ifearulundu and Kaneene, 1998; 1999; Wells and Wagner, 2000; Muskens et al., 2003; Hirst et al., 2004; Ward and Perez, 2004; Aly and Thurmond, 2005; Ridge et al., 2005; Kovich et al., 2006; Scott et al., 2006; Nielsen and Toft, 2007 XII; Kobayashi et al., 2008; Ansari-Lari et al., 2008; Diéguez et al., 2008; Nielsen et al., 2008 XI; Tiwari et al., 2009), but only a limited number of the theoretical risk factors related to management have been confirmed.

Table 4.1. Summary of reported risk factors based on herd-level recording of different outcomes related to infection with *Mycobacterium avium* subsp. *paratuberculosis*

Case definition	Diagnostics	n	Risk factor identified	OR (95% CI)	Reference
Risk of herd having MAP-affected animals					
Clinical disease	Farmer's observation	2470?	Never offering hay to calves	3.3 (1.1; 10)[#]	Cetinkaya et al., 1997
Clinical disease	Farmer's observation	2470?	No use of purchased coarse mixed concentrates	3.3 (1.1; ∞)[#]	Cetinkaya et al., 1997
Clinical disease	Farmer's observation	2470?	No calving in individual pens in grassing periods	4.8 (1.1; ∞)[#]	Cetinkaya et al., 1997
Risk of MAP infection in herd associated with management of around calving and in the calving area					
>0 ELISA+	ELISA	85	No cleaning of maternity pen after each use	3.6 (1.1; 13)	Johnson-Ifearulundu and Kaneene, 1998
>0 ELISA+	ELISA	85	Washing of cows' udders prior to parturition	8.7 (1.9; 40)	Johnson-Ifearulundu and Kaneene, 1998
> 1 ELISA+ or >0 ELISA+ & >5% clinical in past year	ELISA & observation by farmer	963	Use of group housing for periparturient cows during previous year	1.5 (1.0; 2.3)	Wells and Wagner, 2000
>1 ELISA+ cow	ELISA	101	Use of colostrum from cows with positive diagnosis	87 (16; 484)	Diéguez et al. 2008
>1 ELISA+ cow	ELISA	315	More than 1 cow in maternity pen	1.7 (1.2; 2.2)	Tiwari et al., 2009
Risk of MAP infection in herd associated with management of calves					
>0 FC+	FC	33	Exposure of calves 0–6 weeks of age to faeces of adult cows	31 (1.2; 800)	Obasanjo et al., 1997
>0 ELISA+	ELISA	85	No cleaning of calf hutches or pens after each use	3.6 (1.1; 13)	Johnson-Ifearulundu and Kaneene, 1998
> 1 ELISA+ or >0 ELISA+ & >5% clinical in past year	ELISA & observation by farmer	963	Use of group housing for calves prior to weaning previous year	1.5 (1.0; 2.3)	Wells and Wagner, 2000
>1 ELISA+ cow	ELISA	101	Housing of calves <6 months of age with adults	4.6 (1.2; 18)	Diéguez et al. 2008
>1 ELISA+ cow	ELISA	315	Group housing of preweaned calves in winter	2.0 (1.3; 2.8)	Tiwari et al., 2009
Risk of MAP infection in herd associated with other management factors					
>0 ELISA+	ELISA	85	Use of exercise lot for lactating cows	3.0 (1.0; 8.8)	Johnson-Ifearulundu and Kaneene, 1998
> 1 ELISA+ or >0 ELISA+ & >5% clinical in past year	ELISA & observation by farmer	963	Herd size ≥ 300 vs. <50 cows	4.6 (2.3; 9.2)	Wells and Wagner, 2000
>0 ELISA+	ELISA	370	"Large" herd size	?	Muskens et al., 2003
>0 ELISA+	ELISA	85	No application of lime to pastures	10 (1.8; 50)	Johnson-Ifearulundu and Kaneene, 1998
>1 ELISA+ cow	ELISA	315	Purchase of open heifers	2.3 (1.2; 3.5)	Tiwari et al., 2009
>1 ELISA + cow	ELISA	315	Nose-to-nose contact with beef cattle	1.9 (1.1; 2.8)	Tiwari et al., 2009
>1 ELISA+ cow	ELISA	315	Area for pasture <40 hectares vs. > 81 hectares	1.7 (1.1; 3.3)	Tiwari et al., 2009
Risk of occurrence of MAP in bulk tank milk					
PCR+	IS*900* PCR in bulk tank milk	110	Faecal contamination of udders of periparturient cows	6.4 (1.3; 31)	Ansari-Lari et al., 2008

[#]Confounding was not reported investigated.
Abbreviations: +: Test positive; 95% CI: 95% confidence interval; FC: faecal culture; n: sample size (no. of herds); OR: odds ratio

Summaries of reported environmental and management-related risk factors operating at herd- and animal-level, respectively, are shown in Tables 4.1 and 4.2. Table 4.1 is organised with three risk strata: 1) Risk of becoming MAP-affected, defined by occurrence of clinical disease in a herd; 2) Risk of becoming MAP-infected, defined by having one or more ELISA or faecal culture positive animals in a herd; and 3) Risk of occurrence of MAP in bulk tank milk. Stratum 2 is further divided into factors related to a) calving and calving area, b) housing of calves, and c) other factors. In Table 4.2, the primary outcome is ELISA positive at animal level, which is interpreted as risk of a cow being infected. This table is therefore not stratified by outcome, but only by strata related to management at calving, calves and other factors. The data presented in the tables suggest that the majority of hypothetical risk factors (Sweeney, 1996) can be confirmed. They have been implemented in risk assessment tools in various control programmes: for example, the Paraplanner in the Netherlands (Franken, 2005), in the USA (Rossiter et al., 1999) and in Denmark (Nielsen and Nielsen, 2005; 2007). These tools may serve primarily to educate farmers; alternatively, they could be used for risk-factor recording in epidemiological studies.

Table 4.2. Summary of reported risk factors based on animal-level recording of different outcomes related to infection with *Mycobacterium avium* subsp. *paratuberculosis*, and recording of risk factors at herd-level

Case	Test	n	Risk factor identified	OR (95% CI)		Reference
Risk of MAP infection associated with management of around calving and in the calving area						
ELISA +	ELISA	1056	Variation in ELISA of daughter attributable to dam	R^2=7.7%		Nielsen et al., 2002 X
ELISA +	ELISA	510	Born to ELISA positive dam	6.6	(2.2; 20)	Aly and Thurmond, 2005
ELISA +	ELISA	7410	Low amount of straw in bedding in calving area	~3	(1.2; 8.1)	Nielsen and Toft, 2007 XII
ELISA +	ELISA	93994	Feeding colostrum from multiple cows vs. feeding from own dam	1.2	(1.1; 1.4)	Nielsen et al., 2008 XI
Risk of MAP infection associated with management of calves and young stock						
ELISA +	ELISA	510	Exposure as calf to flush water containing adults faeces	28	(8.6; 94)	Aly and Thurmond, 2005
ELISA +	ELISA	93994	Suckling with foster cow vs. feed milk replacer	2.0	(1.4; 3.0)	Nielsen et al., 2008 XI
ELISA +	ELISA	7410	Housing calves 2–12 months in littered common pens or bed stalls with slatted floors vs. tie stalls	~3	(1; 9)	Nielsen and Toft, 2007 XII
ELISA +	ELISA	7410	Housing young stock in littered common pens or bed stalls with slatted floors vs. tie stalls	~3	(1; 9)	Nielsen and Toft, 2007 XII
ELISA +	ELISA	7410	High animal density among young stock	~2.0	(1.0; 4.4)	Nielsen and Toft, 2007 XII
Risk of MAP infection associated with other management factors and environmental factors						
ELISA +	ELISA	236	Not born in VJDSHP herd	8.7	(1.2; 63)	Kovich et al., 2006
FC +	FC	236	Not born in VJDSHP herd	2.3	(1.0; 5.7)	Kovich et al., 2006
ELISA +	ELISA	2819	Herd size ≥ 90 cows	1.8	(1.0; 3.3)	Scott et al., 2006
ELISA +	ELISA	2819	Humid climate	1.9	(1.1; 3.6)	Scott et al., 2006
ELISA +	ELISA	2819	Soil pH <7.0	1.9	(1.1; 3.2)	Scott et al., 2006

Abbreviations: +: Test positive; 95% CI: 95% confidence interval; FC: faecal culture; n: sample size (no. of cows); OR: odds ratio; VJDSHP: Voluntary Johne's Disease Herd Status Program in the USA.

Weaknesses in risk factors studies

The majority of risk factor studies have been cross-sectional. Only a few have been retrospective (Wells and Wagner, 2000; Aly and Thurmond, 2005; Kovich et al., 2006; Nielsen et al., 2008 XI) or prospective (Nielsen and Toft, 2007 XII), and generally much uncertainty about the recording of the risk factors persists. A major strength of the recent study conducted by Nielsen et al. (2008 XI) was a large sample size of almost 100 000 study objects. The fact that only a few risk factors (the feeding of calves with colostrum and milk) were assessed in this retrospective study contributed, in one way, to its strength. It was also a major weakness, however, because it meant that confounding could not be assessed. In Nielsen and Toft (2007 XII) a prospective study was carried out in which risk factors were recorded 26 months prior to the recording of effects determined by ELISA-positivity. A mixture model was used to improve the precision with which animals were categorised as MAP-infected and non-infected. In view of the chronic nature of infection, risk factor studies should be longitudinal, but study designs should be developed to address the issues related to slow development of MAP infections more effectively.

Most studies have been carried out with ELISA or faecal culture (at herd or animal level) as the dependent factor. These tests are inaccurate for assessment of the MAP infection status, a feature that has an impact on the ease with which risk factors can be identified. Furthermore, inaccurate recording of risk factors can affect risk factor identification. The most frequently identified risk factor is “herd size”, probably because this factor is likely to be easy to measure with reasonable precision. Many management-related factors are likely to vary greatly from one animal to another and can therefore be difficult to measure with precision. For example, a calf born at night may receive less attention and be left longer with the dam than a calf born during the day. Unless records of the specific circumstances are made at calving, the calf born during the day and the calf born during night may be recorded as having the same risk. Calves born earlier than expected may be born in the dry-cow area containing several cows – an area that presents a higher risk than a newly cleaned calving area. The recorder of risk factors may only be told that calvings have occurred in newly cleaned calving areas. Furthermore, some management factors may change over time. Inaccurate recording of risk factors, combined with misclassification of infection status, in studies carried out in cross-sectional designs may explain some of the difficulties involved in the identification of risk factors. There is still a great need for studies on management-related risk factors. Models including a more precise dependent variable (Nielsen and Toft, 2007 XII), or bigger sample sizes (Nielsen et al., 2008 XI) may be one solution.

5. DESIGN OF THE CONTROL PROGRAMME

Requirements for establishment of a MAP control programme in Denmark

There was no centrally organised MAP control programme in Denmark until 2006, when a programme was introduced aimed at: 1) providing farmers with tools to manage MAP infections; and 2) reducing the national prevalence of MAP (Nielsen et al., 2007a). The programme is voluntary, and in support of its goals communication and education are considered greatly important. Because certification has not, to date, been specified as an aim, the focus is on implementing control options.

The Danish programme was designed to address the following requirements:

1) Tools for effectively breaking transmission should be available (precondition P1 (Table 1.2)), and these tools should be such as to create minimal additional workload to farmers;
2) Diagnostic testing should be inexpensive and require as little work as possible for farmers (partly P2);
3) Farm-level estimation of production losses should be readily available to justify the establishment of a programme in a specific farm (partly P3);
4) Farmers should be made aware of the limitations of diagnostic tests: for example, the term “free of MAP” should be avoided until a test-strategy to support such a certificate becomes available (P2 and P4);
5) The programme – including resources for programme administration, which would be carried out by the Danish Cattle Federation – should be fully financed by farmers (P5)
6) The programme should be cost-effective (P3).

Design of the Danish control programme

A risk-based approach was deemed an appropriate way to reduce the additional workload required of farmers in managing MAP infections. A test-and-cull strategy would have resulted in the least work, but it would not have been cost-effective (Kudahl et al., 2007). High-risk cows could be identified by frequent testing using a milk antibody ELISA, because this method would create little extra work for the farmer and testing would be relatively cheap. This risk-based approach has many similarities with the approach described by Bang (1908) for managing *M. bovis* infections and is inspired in part on the Danish tuberculosis control scheme.

Repeated testing

The main objective of testing is the detection of animals shedding MAP in doses high enough to result in an infection in a susceptible animal. Although ELISA is generally a relatively sensitive method of detecting animals shedding MAP, frequent testing is considered important because it continuously provides updated test results and hence animal

classifications. For example, approximately 60–70% of high shedders are detected using the ELISA described in Nielsen (2002) if the animal is tested within 3 months prior to the onset of shedding, whereas only approximately 30% are detected if the animal is tested 1 year before shedding starts (Nielsen, 2008 V). Use of the Pourquier ELISA (IDEXX, Montpellier, France) would have resulted in detection of 70–90% of the high shedders prior to shedding with 4 annual tests, and only 25% if the animal was tested one year before shedding (Fig. 2.2). These numbers are point-estimates: they do not take into account any cumulative aspects that might arise from serial or parallel interpretation of the test results.

Management of high-risk animals

Once high-risk animals are identified, it is important that transmission to susceptible animals does not occur. Calves, because they are generally considered most susceptible, require the most rigorous protection. Transmission of MAP occurs via faeces, milk, colostrum and in the uterus. Therefore, calves should at all times be protected from faeces, milk and colostrum obtained from high-risk animals. Moreover, high-risk animals with the highest risk of being infectious should be culled prior to calving. If a system is established to manage these high-risk cows, there is no need to establish routines to manage all cows as if they were infectious. Many farmers may be reluctant to establish such procedures, as they can be very time-consuming. Risk-based management using this system can be cost-effective, but whether a risk-based approach or a non-risk-based approach is most cost-effective depends on the amount of labour required to reduce transmission (Kudahl et al., 2008 XIII). If a heavy workload of management activities intended to reduce MAP transmission is required, a risk-based approach is most cost-effective. Where the workload is light, the test-costs exceed what can be gained from separating cows into high-risk and low-risk groups. Lack of test results also reduces the amount of information available to monitor the test-prevalence over time.

Management of non-affected test-positive animals

A challenge is presented by the management of repeated results, particularly if reacting cows do not appear to be affected by the MAP infection. These animals may or may not be infectious, because antibodies occur in approximately 20% of animals 3–4 years prior to onset of detectable shedding, and they may become test-negative in the meantime (Nielsen, 2008 V). In many cases attempts to confirm that a cow is infected using faecal culture may fail (Nielsen, 2008 V), and it is therefore of the utmost importance to trust the results, even though there is a high risk that false-positive reactions have occurred. A false-positive in this situation may be the result of a) non-MAP antibodies, or b) MAP antibodies in an infected cow that has not yet become infectious. It is not possible to easily discriminate between the two kinds of case. However, a concurrent drop in milk production (Nielsen et al., 2009 VI) might indicate that an infection associated with MAP is ongoing and should result in culling.

Cows with repeated positive ELISA tests have a higher probability of shedding MAP than cows with fluctuating responses (Nielsen, 2008 V), and may be considered for culling. Where there are high prevalences, cows with high ELISA readings should be culled first, as these have the highest probability of shedding MAP (Toft et al., 2005 III). Reactors that are not obviously affected or have fluctuating ELISA responses may be considered of intermediate risk. However, even intermediate risk is an unacceptable risk in terms of transmission if the Bang-method is followed. Therefore, all reacting cows are considered possible sources of MAP, but only those with the highest risk should be culled. Consequently, there are three risk groups. For communication purposes these are referred to as Red (High-Risk, to be culled), Yellow (High-Risk, to keep) and Green (Low-Risk). Management of cows in the three groups can follow the guidance shown schematically in Fig. 5.1.

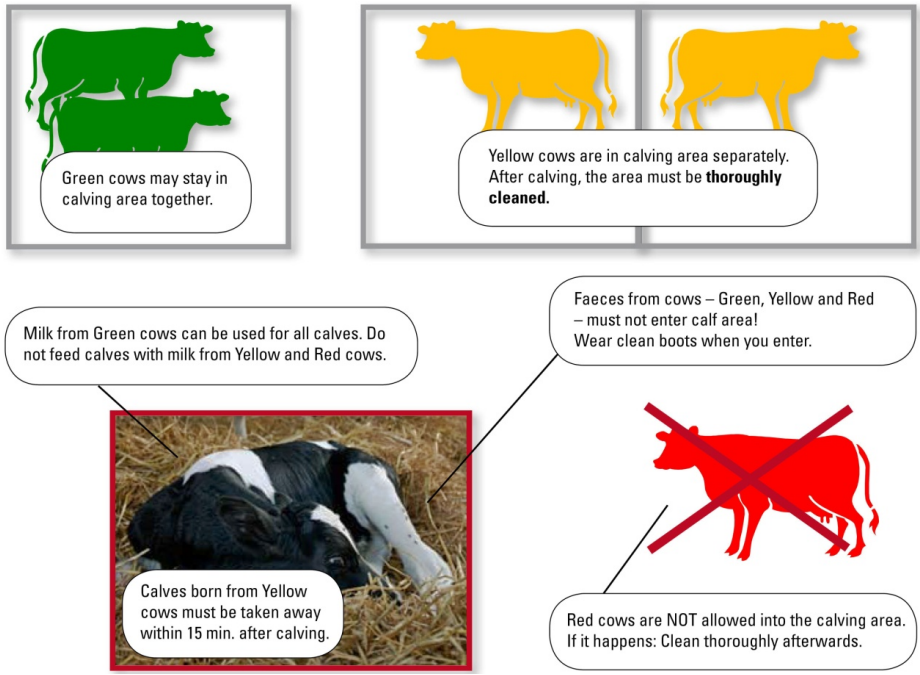


Fig. 5.1. Extract from a poster from the Danish Cattle Federation summarising the principles of the risk-based approach to reducing MAP transmission

Discussion of the Danish programme

The Danish control programme is based on some important assumptions, namely that calves are the most susceptible animals, and that cows are the most infectious animals. These assumptions lead naturally to the conclusion that the main risk factor is calves' exposure to the faeces of infectious cows, but in fact there is little information to confirm them. To date, the long and variable incubation period has been a challenging feature in the study of risk factors, but the data collected in the control programme may be useful in future studies of

risk factors if management factors are also recorded early in the programme. One concern is that such studies may not address is the role of potentially infectious calves (van Roermund et al., 2007). Another centres on farmers' willingness to implement the required changes in management. Implementing these changes may be a major challenge (Ridge et al., 2005), and communication and education is therefore vitally important. These issues are not examined in this thesis.

6. CONCLUSIONS

MAP infections have been known to occur for more than a century, but the work summarised in this thesis and the accompanying papers have added to the preconditions of control and potential eradication. Although the present state of knowledge is far from complete, there is enough solid evidence to establish efforts to control MAP infections. Bovine tuberculosis was controlled with less knowledge than is currently at hand for MAP infections.

Diagnosis

Used and interpreted with care, milk ELISA can be a useful tool in controlling MAP infections. Tests carefully designed according to specific to purposes are essential, particularly because the pathogenesis and incubation period for MAP infections vary greatly. The variation in reported test-accuracies is huge. This is partly due to poor study designs and the lack of a perfect reference test, but stratification into target conditions such as non-infected, infected, infectious and affected ease the interpretation of the test-information (Nielsen and Toft, 2008 I). ELISA detects antibodies, and there is no direct link to any of the target conditions that are relevant to decision makers such as farmers and veterinarians. Therefore, this link needs to be described. The majority of animals that shed detectable amounts of MAP were found to have detectable quantities of antibodies at some point during their lives. The highest probability of testing positive was at 2.5–5.5 years of age (Nielsen and Ersbøll, 2006 II). The probability of MAP shedding, either continuous or intermittent, increased with rising ELISA-level, and rather than dichotomising the ELISA-value the continuous test result provided added value where interpretation of the test-information was concerned (Toft et al., 2005 III). Furthermore, the results presented here suggest that whereas the probability of detecting MAP-infected animals increases with age, the probability of detecting infectious animals is less affected by age (Nielsen and Toft, 2006 IV).

If ELISA is used to detect infectious animals, it is essential to know when antibodies occur relative to MAP shedding. MAP antibodies generally occurred prior to detection of MAP shedding by faecal culture, and in approximately 20% of the animals classified as high shedders antibodies were detected 3–4 years before shedding was detected. If animals were tested on the day of shedding, ~70% would be detected as shedders with the Danish ELISA, whereas if they were tested 1 year before shedding, only 30% would be detected. Therefore, frequent testing is a prerequisite of continuous classification of animals prior to their excretion of MAP (Nielsen, 2008 V). The cumulative probability of true-positives is likely to be higher, but a high test-frequency also results in an increased cumulative frequency of false-positive reactions. However, the grouping of cows on the basis of their repeated test results can be used to draw inferences about MAP shedding and affection due to MAP infection. Cows with a last sample, or the last two samples, positive in ELISA will have a higher probability of shedding MAP on the day of testing; and, by contrast, cows with a fluctuating ELISA response will have a lower probability of shedding MAP (Nielsen, 2008 V). Milk production in

cows with fluctuating immune responses does not appear to be affected by MAP infection. Indeed such cows may actually produce more milk than their herd-mates in same parities and stages of lactation (Nielsen et al., 2009 VI). Only cows with a last positive, or last two positive, ELISA results experience a drop in milk production, but this drop may already occur in some cows 300 days before they test positive in ELISA, and these cows appear to be very likely to be culled due to milk production losses before the necessary diagnosis is confirmed with another ELISA test. In summary, then, the data suggest that MAP-infectious cows can generally be detected using ELISA if frequent testing is carried out, but some MAP-affected cows may be detected primarily because they are affected by MAP infection. One positive ELISA test, combined with the observed drop in milk yield, may provide a good basis for culling of cows in this last category. The results presented here are largely from longitudinal studies, which are currently sparse in literature. Further longitudinal studies confirming the results presented here will ideally be conducted by other researchers. If test results are to be used to drive inferences about the MAP-infected status of the individual, ELISA used for detecting MAP-infected animals needs further characterisation.

Prevalence

Absolute prevalence estimates of MAP infections are poor (Nielsen and Toft, 2009 VII). This is partly because reported sensitivities of ELISA for the detection of MAP infection are highly variable (Nielsen and Toft, 2008 I) and therefore true prevalences cannot be estimated. Furthermore, many prevalence studies were conducted on populations not reflecting the target conditions. However, the available data do suggest that prevalences in many European cattle populations are high, and data from multiple studies in Denmark, conducted at both herd and animal level, also suggest that the prevalence is high, although the absolute prevalence cannot be estimated with complete certainty (Nielsen and Toft, 2009 VII). Studies that most reliably reflect absolute prevalence may best be conducted using tissue culture and histological data obtained from abattoir cattle, as was done by McKenna et al. (2004). However, such studies are expensive and impractical and may include an age bias. The sample size obtainable in such a study may be insufficient to permit investigation of risk factors, and the sampling frame will often have to be cross-sectional to be feasibly carried out in practice. For certification purposes the approach may be too expensive. Alternatives are ELISA, eventually with follow-up confirmation using faecal culture, or faecal culture alone. In the Danish cattle population, with an expected high prevalence of MAP-infected herds and animals, whole-herd ELISA of all lactating animals appears to be the most cost-effective method to estimate the probability that a herd has a low prevalence (Sergeant et al., 2008 VIII). Estimated median probabilities for most herds were >0.9, although larger herds (>80 cows) will be easier to classify than smaller herds. Correction for age-corrected sensitivities can also facilitate the provision of prevalence estimates, but such estimates are required for each specific ELISA and will not necessarily be transferrable to other ELISAs.

Age-stratified sensitivity and specificity estimates are often not available, and where they are available they may be associated with a high degree of uncertainty. A Bayesian mixture model using the continuous ELISA response was developed to estimate within-herd prevalences of MAP infection (Nielsen et al., 2007b IX). Parity and days in milk were included as approximations of age. This model provided more precise prevalence estimates than cut-off based methods. However, true prevalences were not obtained with this model either, and so challenges in providing unbiased prevalence estimates of MAP infection within herds persist. Prevalence estimates should be considered relative prevalences. The latter can only be compared with estimates using the same test.

Risk factors

Multiple risk-factor studies of the transmission of MAP have been carried out, but few risk factors have been confirmed or identified. A relatively specific measure, such as being born to a specific dam (Nielsen et al., 2002 X), has been demonstrated as a risk factor, suggesting that it is possible to demonstrate these factors with existing tests. However, more precise prevalence-ranking of herds, in combination with retrospective or prospective studies, may be required in studies of some of the management-related factors, since such factors are subject to high variation within a herd. A retrospective study design was used to assess the effect of source of milk and colostrum fed to calves as risk factors for becoming ELISA positive as cows (Nielsen et al., 2008 XI). Both milk and colostrum were identified as significant risk factors. The specific factors resulting in increased odds of infection: calves fed colostrum from multiple dams compared to receiving colostrum exclusively from their own dams (OR=1.24, 95% CI: 1.09; 1.42); and letting calves suckle with foster cows (OR=2.01; 95% CI: 1.37; 2.96). The latter may simply be due to the calf's staying in an environment where, possibly, there is high faecal contamination from cows. The estimates for sources of milk and colostrum were not corrected for the potentially biggest confounder, namely factors that would have an impact on exposure to the faeces of adult cattle. A prospective study using the Bayesian mixture model (Nielsen et al., 2007b IX) revealed risk factors related to management that had not previously been identified: calves and young stock housed in littered common pens, or bed stalls with slatted floors, had 3 times higher odds of being sero-positive than animals raised in tie stalls; and a small amount of straw in the bedding of the calving area and high animal density among young stock also raised the odds of animals being found to be sero-positive (Nielsen and Toft, 2007 XII). Further exploration of management-related risk factors could be undertaken using the Bayesian mixture model.

Establishment of the control programme

The detection of MAP-infected animals, prevalence assessment and risk factor studies are all challenging. However, calves' contact with the faeces of infectious animals is the primary management factor that needs to be addressed in efforts to control MAP infections. If the

infectious cows can be detected before this contact, risk-based management is feasible. It was demonstrated that if cows are tested close to the start of bacterial shedding, ELISA has a reasonably high sensitivity (Nielsen, 2008 V). Simulation studies were then used to assess whether a risk-based approach with four annual herd tests would be both cost-effective and, simultaneously, such as to reduce the workload required by management of high-risk animals (Kudahl et al., 2008 XIII). The risk-based approach was more cost-effective than a non-risk-based approach if changes in management required more than 1 hour per calving. However, in the non-risk-based situation, there would be no test results available to monitor developments in prevalence.

7. FUTURE PERSPECTIVES

Of the approximately 4 430 Danish dairy herds, 1 250 (28%) had participated in the control programme by January 2009. Reasons for lack of participation are not known, but it is possible that lack of a certification scheme is a major consideration for some herd owners. Farmers with herds free of animals affected by MAP would logically chose certification programmes – at any rate, if such an option were presented, and if they could see the benefit of it. There are several options for establishing a certification scheme, but none of the existing test-strategies is capable of providing high probabilities that herds are free of MAP infection. On the other hand, Sergeant et al. (2008 VIII) have suggested that whole-herd ELISA without confirmatory testing would provide reasonably high probabilities that herds have a low prevalence of MAP. Moreover, whole-herd ELISA was the most cost-effective approach when compared with whole-herd faecal culture, ELISA on only a part of the animals, and ELISA with confirmatory testing using faecal culture, particularly for herds > 80 cows. Use of ELISA results would have the dual advantage that only one test-system would be needed, and that test results could be used for management in infected herds. Use of historical data may be a useful supplement to the model suggested in Sergeant et al. (2008 VIII), but this approach needs to be sensitive to correlations between test results from different periods; it must also be able to take into account the introduction of animals into a herd.

The fact that more than 25% of the farms in Denmark have joined the programme suggests that some of its tools are considered acceptable by a substantial number of farmers, despite testing costs being >10 EUR per cow per year. Currently no measures to determine whether changes in management occur, and whether these changes have an effect. However, failure to make the relevant changes will result in a corresponding failure to control MAP infections, and the costs for testing will be wasted. Also, as it is the next generation that is supposed to be free of MAP, and infected animals among this generation will only affect prevalence when they become adult, it takes a while before an effect, or lack thereof, becomes visible in the results. Therefore, there is a need for a monitoring tool to detect early changes in management, and to determine whether calves are still being exposed to MAP. Because cell-mediated immune reactions are the first to appear, a diagnostic test to be applied to young animals would be desirable. Young animals found to be positive in a cell-mediated test may be able to control and eradicate their own MAP infection and should therefore not be culled, especially as these animals may have valuable genetic traits. However, if they are test-positive, and the test result is specific, a positive result could indicate that they have been exposed to MAP (which they should not have been if control measures were sufficient to break transmission).

So long as management factors are recorded in all herds early in the process, the herds and animals in the control programme can serve as a valuable means of studying further risk factors. Such recordings may still be inaccurate, although prospective studies are

likely to be less inaccurate than retrospective studies. However, there is still a great need to understand management-related risk factors in order to further optimise risk-based management. For example, milk used for feeding of calves appears not to be an important risk factor, but scenarios in which numerous MAP infections can result are likely to occur if, for example, milk from MAP-affected cows is fed to calves. Delivering such milk to the dairy is not allowed, because only the milk of healthy cows may be so delivered for human consumption (Anon., 2004b, pp. 119–130). Therefore, some farmers may feed high concentrations of MAP to some animals in certain unfortunate circumstances, and this may not be properly reflected in most data. Furthermore, because animals are continuously monitored over time, it may be possible to obtain more information about the incubation period and thereby determine whether adult animals should be afforded greater concern as susceptible animals.

If eradication is to be achieved regionally in Denmark, there is still a need for a source of low-risk replacement animals. However, there is also a need to demonstrate that control can be achieved in real herds, and that these herds can subsequently be certified as free of MAP infection. These steps are still to be taken. Tools to monitor “success” are urgently needed, because farmers with a long-term control programme ahead need continuous and renewable motivation to maintain good management practices. The results of the control and eradication of *M. bovis* demonstrated that individual herds could control the infection, but assistance from industry-ownership, and suitable regulatory systems, was required to eradicate the infection from entire regions. Today, however, the number of herds is smaller, and the testing of animals is easier if milk ELISA is employed. Moreover, although the incubation period for MAP appears to be longer than it is for *M. bovis* infections, it ought to be feasible to control MAP using the principles described by Bang more than a century ago. Herd size has increased significantly over the years, certainly, but Bang actually established his programmes also in herds of more than 100 cows.

REFERENCES

Aly SS, Thurmond MC, 2005. Evaluation of *Mycobacterium avium* subsp *paratuberculosis* infection of dairy cows attributable to infection status of the dam. J Am Vet Med Assoc. 227: 450–454.

Andersen HJ, Aagaard K, Skjøth F, Rattenborg E, Enevoldsen C, 2000. Integration of research, development, health promotion, and milk quality assurance in the Danish dairy industry. *In*: Salman MD, Morley P, Ruch-Gallie R (Eds): Proceedings of the 9th symposium of the International Society of Veterinary Epidemiology and Economics, Breckenridge, Colorado, USA, Aug 6–11, pp. 258–260.

Andersen S, 1965. Paratuberkulose hos kvæg. En oversigt over sygdommens kliniske manifeste optræden i Danmark (In Danish). Medlemsblad Den Danske Dyrlægeforening, 48: 409–414.

Andrews JM, Langmuir AD, 1963. The philosophy of disease eradication. Am J Public Health Nations Health. 53: 1–6.

Anon., 1932. Lov nr. 196 af 26. juni 1932. Lov om bekæmpelse af tuberkulose hos kvæg og svin (In Danish). Danish Ministry of Agriculture. Lovtidende A. nr. 38, p. 1177–1180.

Anon., 2003. Validation and certification of diagnostic assays. Resolution no. XXIX by the International Committee to the OIE. http://www.oie.int/vcda/eng/en_background_vcda.htm (accessed Dec. 27, 2008).

Anon., 2004a. The risk of transmission of *Mycobacterium avium* subsp. *paratuberculosis* via bovine semen. EFSA J: 110, 1–59.

Anon., 2004b. Regulation (EC) No 853/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific hygiene rules for food of animal origin. Official J EU, L139: 55.

Ansari-Lari M, Haghkhah M, Bahramy A, Novin Baheran AM, 2008. Risk factors for *Mycobacterium avium* subspecies *paratuberculosis* in Fars province (Southern Iran) dairy herds. Trop Anim Health Prod (in press). doi: 10.1007/s11250-008-9221-7

Antognoli MC, Garry FB, Hirst HL, Lombard JE, Dennis MM, Gould DH, Salman MD, 2008. Characterization of *Mycobacterium avium* subspecies *paratuberculosis* disseminated infection in dairy cattle and its association with antemortem test results. Vet Microbiol. 127: 300-308.

Bang B, 1892. Om tuberkulinet og det anvendelse i kampen mod kvægtuberkulosen (In Danish). Ugeskrift for Landmænd, 7: 1–4 & 15–22.

Bang B, 1906. Chronische pseudotuberkulöse – Darmentzündung beim Rinde (In German). Berliner Tierärztlich Wochenschrift, 26: 759–763.

Bang B, 1908. The Bang Method for the repression of tuberculosis in cattle. Commonwealth of Pennsylvania, Department of Agriculture, Bulletin no. 172, Harrisburg, Pennsylvania, USA.

Bang B, 1909. Kvægets smitsomme kroniske Tarmbetændelse (In Danish). Maanedsskrift for Dyrlæger, 21: 513–548.

Bang B, 1928. Tuberculosis in cattle. J Am Vet Med Assoc. 72: 20–25.

Baptista FM, Nielsen SS, Toft N, 2008. Association between presence of antibodies to *Mycobacterium avium* subsp. *paratuberculosis* and somatic cell count in dairy cows. J. Dairy Sci., 91: 109–118.

Benedictus G, Dijkhuizen AA, Stelwagen J, 1987. Economic losses due to paratuberculosis in dairy cattle. Vet Rec. 121:142–146.

Benedictus G, Verhoeff J, Schukken YH, Hesselink JW, 2000. Dutch paratuberculosis programme history, principles and development. Vet Microbiol. 77: 399–413.

Bielanski A, Algire J, Randall GC, Surujballi O, 2006. Risk of transmission of *Mycobacterium avium* ssp. *paratuberculosis* by embryo transfer of in vivo and in vitro fertilized bovine embryos. Theriogenology. 66: 260–266.

Bosshard C, Stephan R, Tasara T, 2006. Application of an F57 sequence-based real-time PCR assay for *Mycobacterium paratuberculosis* detection in bulk tank raw milk and slaughtered healthy dairy cows. J Food Prot. 69: 1662–1667.

Buergelt CD, Williams JE, 2004. Nested PCR on blood and milk for the detection of *Mycobacterium avium* subsp *paratuberculosis* DNA in clinical and subclinical bovine paratuberculosis. Aust Vet J. 82: 497–503.

Buergelt CD, Hall C, McEntee K, Duncan JR, 1978. Pathological evaluation of paratuberculosis in naturally infected cattle. Vet Pathol. 15: 196–207.

Bulaga LL, 1998. U.S. Voluntary Johne's Disease Program for cattle. Proceedings of the 102nd Annual meeting of the United Status Animal Health Association, Minneapolis, Minnesota, USA, October 3–9, 1998, p. 420–433.

Buza JJ, Hikono H, Mori Y, Nagata R, Hirayama S, Aodon-geril, Bari AM, Shu Y, Tsuji NM, Momotani E, 2004. Neutralization of interleukin-10 significantly enhances gamma interferon expression in peripheral blood by stimulation with Johnin purified protein derivative and by infection with *Mycobacterium avium* subsp. *paratuberculosis* in experimentally infected cattle with paratuberculosis. Infect Immun. 72: 2425–2428.

Böttcher J, Gangl A, 2005. Value of bulk-milk serology for control of Johne's disease. Proceedings of the 8th International Colloquium on Paratuberculosis, Aug. 14–18, 2005, Copenhagen, Denmark, p. 477–484.

Carter MA, 2007. An overview of the Voluntary Bovine Johne's Disease Control Program in the United States of America. Bulletin of IDF. 410: 14–19.

Cetinkaya B, Erdogan HM, Morgan KL, 1997. Relationships between the presence of Johne's disease and farm and management factors in dairy cattle in England. Prev Vet Med. 32: 253–266.

Christiansen MJ, 1935. Erfaringer og resultater ved tuberkulosebekæmpelsen paa isoleret beliggende øer og i mejerikredse efter §5 i Lov af 26. juni 1932 samt om de aviære

infektioners udbredelse blandt kvæget på disse øer (In Danish). Maanedsskrift for Dyrlæger, 47: 321–336.

Christiansen MJ, 1941. Kvægtuberkulosens bekæmpelse i Jylland (In Danish). Jydsk Landbrug, 23: 237–241.

Clark DL Jr, Koziczkowski JJ, Radcliff RP, Carlson RA, Ellingson JL, 2008. Detection of *Mycobacterium avium* subspecies *paratuberculosis*: comparing fecal culture versus serum enzyme-linked immunosorbent assay and direct fecal polymerase chain reaction. J Dairy Sci. 91: 2620–2627.

Collins DM, Gabric DM, De Lisle GW, 1989. Identification of a repetitive DNA sequence specific to *Mycobacterium paratuberculosis*. FEMS Microbiol Lett. 51: 175–178.

Collins DM, Cavaignac S, de Lisle GW, 1997. Use of four DNA insertion sequences to characterize strains of the *Mycobacterium avium* complex isolated from animals. Mol Cell Probes. 11:373–380.

Collins MT, 1996. Diagnosis of paratuberculosis. Vet Clin North Am Food Anim Pract. 12: 357–371.

Collins MT, 2002. Interpretation of a commercial bovine paratuberculosis enzyme-linked immunosorbent assay by using likelihood ratios. Clin Diagn Lab Immunol. 9: 1367–1371.

Collins MT, Sockett DC, Goodger WJ, Conrad TA, Thomas CB, Carr DJ, 1994. Herd prevalence and geographic distribution of, and risk factors for, bovine paratuberculosis in Wisconsin. J Am Vet Med Assoc. 204: 636–641.

Collins P, Matthews PRJ, McDiarmid A, Brown A, 1983. The pathogenicity of *Mycobacterium avium* and related mycobacteria for experimental animals. J Med Microbiol, 16, 27–35.

Collins P, McDiarmid A, Thomas LH, Matthews PR, 1985. Comparison of the pathogenicity of *Mycobacterium paratuberculosis* and *Mycobacterium* spp isolated from the wood pigeon (*Columba palumbus L*). J Comp Pathol. 95: 591–597.

Coussens PM, 2001. *Mycobacterium paratuberculosis* and the bovine immune system. Anim Health Res Rev. 2: 141–161.

Coussens PM, 2004. Model for immune responses to *Mycobacterium avium* subspecies *paratuberculosis* in cattle. Infect Immun, 72: 3089–3096.

de Almeida DE, Colvin CJ, Coussens PM, 2008. Antigen-specific regulatory T cells in bovine paratuberculosis. Vet Immunol Immunopathol. 125: 234–245.

Diéguez FJ, Arnaiz I, Sanjuán ML, Vilar MJ, Yus E, 2008. Management practices associated with *Mycobacterium avium* subspecies *paratuberculosis* infection and the effects of the infection on dairy herds. Vet Rec. 162: 614–617.

Doyle TM, 1953. Susceptibility to Johne's disease in relation to age. Vet Rec, 65: 363–365.

Doyle TM, Spears HN, 1951. A Johne's Disease survey. Vet Rec. 63: 355–359.

Dussol C, 2004. Relations between within-herd-prevalence and bulk tank milk responses in an ELISA for detection of antibodies to *Mycobacterium avium* subsp. *paratuberculosis*. Ecole Nationale Veterinaire de Nantes, France, 88 pp.

Eamens GJ, Whittington RJ, Marsh IB, Turner MJ, Saunders V, Kemsley PD, Rayward D, 2000. Comparative sensitivity of various faecal culture methods and ELISA in dairy cattle herds with endemic Johne's disease. Vet Microbiol. 77: 357–367.

Eamens GJ, Whittington RJ, Turner MJ, Austin SL, Fell SA, Marsh IB, 2007. Evaluation of radiometric faecal culture and direct PCR on pooled faeces for detection of *Mycobacterium avium* subsp. *paratuberculosis* in cattle. Vet Microbiol. 125: 22–35.

Franken P, 2005. Paratuberculosis control in the Netherlands: the target and an overview of activities. Proceedings of the 8th International Colloquium on Paratuberculosis, August 14–18, 2005, Copenhagen, Denmark, p. 15–19.

Giese SB, Ahrens P, 2000. Detection of *Mycobacterium avium* subsp. *paratuberculosis* in milk from clinically affected cows by PCR and culture. Vet Microbiol. 77: 291–297.

Gonda MG, Chang YM, Shook GE, Collins MT, Kirkpatrick BW, 2006. Genetic variation of *Mycobacterium avium* ssp. *paratuberculosis* infection in US Holsteins. J Dairy Sci. 89: 1804–1812.

Grant IR, Ball HJ, Rowe MT, 2002. Incidence of *Mycobacterium paratuberculosis* in bulk raw and commercially pasteurized cows' milk from approved dairy processing establishments in the United Kingdom. Appl Environ Microbiol. 68: 2428–2435.

Hagan WA, 1938. Age as a factor in susceptibility to Johne's disease. Cornell Vet. 28: 34–40.

Hansen P, 1930. Aarsberetning fra Veterinærfysikatet for aaret 1929 (In Danish). Kandrup & Wunsch, Copenhagen, Denmark, pp. 34–37.

Hansen P, 1931. Aarsberetning fra Veterinærdirektoratet for aaret 1930 (In Danish). Kandrup & Wunsch, Copenhagen, Denmark, pp. 34–36.

Hansen P, 1932. Aarsberetning fra Veterinærdirektoratet for aaret 1931 (In Danish). Kandrup & Wunsch, Copenhagen, Denmark, pp. 34–36.

Hansen P, 1933. Aarsberetning fra Veterinærdirektoratet for aaret 1932 (In Danish). Kandrup & Wunsch, Copenhagen, Denmark, pp. 37–40.

Hansen P, 1941. Aarsberetning fra Veterinærdirektoratet for aaret 1940 (In Danish). A. Johansens Bogtrykkeri, Copenhagen, Denmark, pp. 36–44.

Hendrick SH, Kelton DF, Leslie KE, Lissemore KD, Archambault M, Duffield TF, 2005. Effect of paratuberculosis on culling, milk production, and milk quality in dairy herds. J Am Vet Med Assoc. 227: 1302–1308.

Herthnek D, Nielsen SS, Lindberg A, Bölske G, 2008. A robust method for bacterial lysis and DNA purification to be used with real-time PCR for detection of *Mycobacterium avium* subsp. *paratuberculosis* in milk. J Microbiol Methods. 75: 335–340.

Hines II ME, Stabel JR, Sweeney RW, Griffin F, Talaat AM, Bakker D, Benedictus G, Davis WC, de Lisle GW, Gardner IA, Juste RA, Kapur V, Koets A, McNair J, Pruitt G, Whitlock

RH, 2007. Experimental challenge models for Johne's disease: a review and proposed international guidelines. Vet Microbiol. 122: 197–222.

Hirsh DC, Biberstein EL, 2004. *Mycobacterium*. *In*: Hirsh DC, MacLachlan NJ, Walker RL (eds.). Veterinary Microbiology, 2nd Edition, Blackwell Publishing, Ames, Iowa, USA, p. 229.

Hirst HL, Garry FB, Morley PS, Salman MD, Dinsmore RP, Wagner BA, McSweeney KD, Goodell GM, 2004. Seroprevalence of *Mycobacterium avium* subsp *paratuberculosis* infection among dairy cows in Colorado and herd-level risk factors for seropositivity. J Am Vet Med Assoc. 225: 97–101.

Houe H, Ersbøll AK, Nielsen LR, 2004. Nature of data. *In*: Houe H, Ersbøll AK, Toft N (eds.). Introduction to Veterinary Epidemiology, Biofolia, Frederiksberg C, Denmark, p. 33.

Hutchinson LJ, 1996. Economic impact of paratuberculosis. Vet Clin North Am Food Anim Pract. 12: 373–381.

Johnson-Ifearulundu YJ, Kaneene JB, 1998. Management-related risk factors for *M. paratuberculosis* infection in Michigan, USA, dairy herds. Prev Vet Med. 37: 41–54.

Johnson-Ifearulundu Y, Kaneene JB, 1999. Distribution and environmental risk factors for paratuberculosis in dairy cattle herds in Michigan. Am J Vet Res. 60: 589–596.

Johnson-Ifearulundu YJ, Kaneene JB, Sprecher DJ, Gardiner JC, Lloyd JW, 2000. The effect of subclinical *Mycobacterium paratuberculosis* infection on days open in Michigan, USA, dairy cows. Prev Vet Med. 46: 171–181.

Jubb TF, Galvin JW, 2004. Effect of a test and control program for bovine Johne's disease in Victorian dairy herds 1992 – 2002. Aust Vet J. 82: 228–232.

Jørgensen JB, 1965. On the occurrence of *Mycobacterium johnei* in the mesenteric lymph nodes of abattoir cattle. Nord Vet Med. 17: 97–102.

Jørgensen JB, 1972. Undersøgelser over forekomst af paratuberkulose hos kvæg i Danmark (In Danish). Nord Vet Med. 24: 297–308.

Jørgensen JB, 1977. Survival of *Mycobacterium paratuberculosis in slurry*. Nord Vet Med. 29: 267–270.

Jørgensen JB, 1978. Serological investigation of strains of *Mycobacterium avium* and *Mycobacterium intracellulare* isolated from animal and non-animal sources. Nord Vet-Med. 30: 155–162.

Kalis CH, Hesselink JW, Barkema HW, Collins MT, 2000. Culture of strategically pooled bovine fecal samples as a method to screen herds for paratuberculosis. J Vet Diagn Invest. 12: 547–551.

Kennedy DJ, 2007. Development in the approach to managing paratuberculosis in Australia. Bulletin of IDF. 410: 8–13.

Kennedy D, Nielsen SS, 2007. Report from the first IDF ParaTB Forum, Bulletin of IDF. 410: 3–7.

Kobayashi S, Tsutsui T, Yamamoto T, Nishiguchi A, 2007. Current surveillance and control program on paratuberculosis in Japan. Bulletin of IDF. 410: 30–33.

Kobayashi S, Tsutsui T, Yamamoto T, Nishiguchi A. 2008. Epidemiologic indicators associated with within-farm spread of Johne's disease in dairy farms in Japan. J Vet Med Sci. 69: 1255–1258.

Kovich DA, Wells SJ, Friendshuh K, 2006. Evaluation of the Voluntary Johne's Disease Herd Status Program as a source of replacement cattle. J Dairy Sci. 89: 3466–3470.

Kudahl A, Nielsen SS, Sørensen JT, 2004. Relationship between antibodies against *Mycobacterium avium* subsp. *paratuberculosis* in milk and shape of lactation curves. Prev Vet Med. 62: 119–134.

Kudahl AB, Østergaard S, Sørensen JT, Nielsen SS, 2007. A stochastic model simulating paratuberculosis in a dairy herd. Prev Vet Med. 78: 97–117.

Kudahl AB, Nielsen SS, Østergaard S, 2008. Economy, efficacy and feasibility risk-based control program against paratuberculosis. J Dairy Sci. 91: 4599–4609.

Kulkas L, 2007. Plans for a paratuberculosis certification and eradication program in Finland. Bulletin of IDF. 410: 20–22.

Larsen AB, Merkal RS, Cutlip RC, 1975. Age of cattle as related to resistance to infection with *Mycobacterium paratuberculosis*. Am J Vet Res, 36: 255–257.

Lombard JE, Garry FB, McCluskey BJ, Wagner BA, 2005. Risk of removal and effects on milk production associated with paratuberculosis status in dairy cows. J Am Vet Med Assoc. 227: 1975–1981.

Lombard JE, Wagner BA, Smith RL, McCluskey BJ, Harris BN, Payeur JB, Garry FB, Salman MD, 2006. Evaluation of environmental sampling and culture to determine *Mycobacterium avium* subspecies *paratuberculosis* distribution and herd infection status on US dairy operations. J Dairy Sci. 89: 4163–4171.

Martin PA, 2008. Current value of historical and ongoing surveillance for disease freedom: surveillance for bovine Johne's disease in Western Australia. Prev Vet Med. 84: 291–309.

Martin SW, Meek AH, Willeberg P, 1987. Veterinary Epidemiology, p. 245–258. Iowa State University Press, Ames, Iowa, USA.

Matthews PR, McDiarmid A, 1979. The production in bovine calves of a disease resembling paratuberculosis with a *Mycobacterium* sp isolated from a woodpigeon (*Columba palumbus L*). Vet Rec. 104: 286.

McKenna SL, Keefe GP, Barkema HW, McClure J, VanLeeuwen JA, Hanna P, Sockett DC, 2004. Cow-level prevalence of paratuberculosis in culled dairy cows in Atlantic Canada and Maine. J Dairy Sci. 87: 3770–3777.

Merkal RS, Larsen AB, Booth GD, 1975. Analysis of the effect of inapparent bovine paratuberculosis. Am J Vet Res. 36: 837–838.

M'Fadyean J, Sheather AL, 1916. Johne's disease. The experimental transmission of the disease to cattle, sheep, and goats, with notes regarding the occurrence of natural cases in sheep and goats. J Comp Pathol, 29, 62–94.

Mijs W, de Haas P, Rossau R, Van der Laan T, Rigouts L, Portaels F, van Soolingen D, 2002. Molecular evidence to support a proposal to reserve the designation *Mycobacterium avium* subsp. *avium* for bird-type isolates and '*M. avium* subsp. *hominissuis*' for the human/porcine type of *M. avium*. Int J Syst Evol Microbiol. 52: 1505–1518.

Momotani E, Whipple DL, Thiermann AB, Cheville NF, 1988. Role of M cells and macrophages in the entrance of *Mycobacterium paratuberculosis* into domes of ileal Peyer's patches in calves. Vet Pathol. 25: 131–137.

Mortensen H, Nielsen SS, Berg P, 2004. Genetic variation and heritability of the antibody response to *Mycobacterium avium* subspecies *paratuberculosis* in Danish Holstein cows. J Dairy Sci. 87: 2108–2113.

Muskens J, Elbers AR, van Weering HJ, Noordhuizen JP, 2003. Herd management practices associated with paratuberculosis seroprevalence in Dutch dairy herds. J Vet Med B. 50: 372–377.

Nielsen FW, 1942. Aarsberetning fra Veterinærdirektoratet for aaret 1941 (In Danish). A. Johansens Bogtrykkeri, Copenhagen, Denmark, pp. 34–41.

Nielsen FW, 1943. Aarsberetning fra Veterinærdirektoratet for aaret 1942 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 29–35.

Nielsen FW, 1944. Aarsberetning fra Veterinærdirektoratet for aaret 1943 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 29–36.

Nielsen FW, 1945. Aarsberetning fra Veterinærdirektoratet for aaret 1944 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 32–38.

Nielsen FW, 1947. Aarsberetning fra Veterinærdirektoratet for aaret 1945 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 28–34.

Nielsen FW, 1948. Aarsberetning fra Veterinærdirektoratet for aaret 1946 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 32–38.

Nielsen FW, 1949a. Aarsberetning fra Veterinærdirektoratet for aaret 1947 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 30–35.

Nielsen FW, 1949b. Aarsberetning fra Veterinærdirektoratet for aaret 1948 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 27–35.

Nielsen FW, 1951. Aarsberetning fra Veterinærdirektoratet for aaret 1949 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 31–38.

Nielsen FW, 1952. Aarsberetning fra Veterinærdirektoratet for aaret 1950 (In Danish). J. H. Schultz A/S Universitets-Bogtrykkeri, Copenhagen, Denmark, pp. 28–32.

Nielsen LR, Nielsen SS, 2005. The use of risk assessment, test and management strategies to control paratuberculosis and *Salmonella* Dublin in dairy herds. Proceedings of the

8th International Colloquium on Paratuberculosis, Aug. 14–18, 2005, Copenhagen, Denmark, pp. 218–225.

Nielsen SS, 2000. Forekomst af paratuberkulose på ko-niveau og besætningsniveau (In Danish). The Newsletter Kongeåprojektet no. 10, Nov. 3, 2000, p. 4.

Nielsen SS, 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subspecies *paratuberculosis* in dairy cattle. J Vet Med B Infect Dis Vet Public Health. 49: 384–387.

Nielsen SS, 2007. Danish control programme for bovine paratuberculosis. Cattle Practice. 15: 161–168.

Nielsen SS, 2008. Transitions in diagnostic tests used for detection of *Mycobacterium avium* subsp. *paratuberculosis* infections in cattle. Vet Microbiol. 132: 274–282.

Nielsen SS, Ersbøll AK, 2006. Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows. J Dairy Sci. 89: 4557–4566.

Nielsen SS, Nielsen LR, 2007. Indsats mod paratuberkulose og *Salmonella* Dublin - rådgivermanual (In Danish). Dansk Landbrugsrådgivning, Århus N, Denmark.

Nielsen SS, Toft N. 2006. Age-specific characteristics of ELISA and fecal culture for purpose specific testing for paratuberculosis. J Dairy Sci. 89: 569–579.

Nielsen SS, Toft N, 2007. Assessment of management-related risk factors for paratuberculosis in Danish dairy herds using Bayesian mixture models. Prev Vet Med. 81: 306–317.

Nielsen SS, Toft N, 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques. Vet Microbiol. 129: 217–235.

Nielsen SS, Toft N, 2009. A review of prevalences of paratuberculosis in farmed animals in Europe. Prev Vet Med. 88: 1–19.

Nielsen SS, Thamsborg SM, Houe H, Bitsch V, 2000. Bulk-tank milk ELISA antibodies for estimating the prevalence of paratuberculosis in Danish dairy herds. Prev Vet Med. 44: 1–7 (with Corrigendum in Prev Vet Med. 46: 297).

Nielsen SS, Nielsen KK, Huda A, Condron R, Collins MT, 2001. Diagnostic techniques for paratuberculosis. Bulletin IDF, 362: 5–17.

Nielsen SS, Gröhn YT, Quaas RL, Agger JF, 2002. Paratuberculosis in dairy cattle: Variation of the antibody-response in offspring attributable to the dam. J Dairy Sci. 85: 406–412.

Nielsen SS, Houe H, Ersbøll AK, Toft N, 2004. Evaluating diagnostic tests. *In*: Houe H, Ersbøll AK, Toft N (Eds): Introduction to Veterinary Epidemiology. Biofolia, Frederiksberg, Denmark, p. 133–151.

Nielsen SS, Enevoldsen C, Toft N, 2006. Milk production losses associated with bovine paratuberculosis diagnosed from repeated testing. Proceedings of the 11th International Symposium on Veterinary Epidemiology and Economics, 6–11 August 2006. Cairns, Australia, p. 619–621.

Nielsen SS, Jepsen ØR, Aagaard K, 2007a. Control programme for paratuberculosis in Denmark. Bulletin of IDF. 410: 23–29.

Nielsen SS, Toft N, Jørgensen E, Bibby BM, 2007b. Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response. Prev Vet Med. 81: 290–305.

Nielsen SS, Bjerre H, Toft N, 2008. Colostrum and milk as risk factors for infection with *Mycobacterium avium* subsp. *paratuberculosis* in dairy cattle. J Dairy Sci. 91: 4610–4615.

Nielsen SS, Krogh MA, Enevoldsen C, 2009. Time to occurrence of a decline in milk production in cows with various paratuberculosis antibody profiles. J Dairy Sci. 92: 149–155.

Obasanjo IO, Gröhn YT, Mohammed HO, 1997. Farm factors associated with the presence of *Mycobacterium paratuberculosis* infection in dairy herds on the New York State Paratuberculosis Control Program. Prev Vet Med. 32: 243–251.

Ott SL, Wells SJ, Wagner BA, 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev Vet Med. 40: 179–192.

Pérez V, Gongález J, Geijo MV, Reyes LE, Garrido M, Corpa JM, García-Pariente C, García Marin JF, 2005. Pathological forms and peripheral immune response in natural bovine paratuberculosis. Proceedings of the 7th International Colloquium on Paratuberculosis, Bilbao, Spain, p. 103–109.

Raizman EA, Wells SJ, Godden SM, Bey RF, Oakes MJ, Bentley DC, Olsen KE, 2004. The distribution of *Mycobacterium avium* ssp. *paratuberculosis* in the environment surrounding Minnesota dairy farms. J Dairy Sci. 87: 2959–2966.

Raizman EA, Fetrow J, Wells SJ, Godden SM, Oakes MJ, Vazquez G, 2007a. The association between *Mycobacterium avium* subsp. *paratuberculosis* fecal shedding or clinical Johne's disease and lactation performance on two Minnesota, USA dairy farms. Prev Vet Med. 78: 179–195.

Raizman EA, Wells SJ, Godden SM, Fetrow J, Oakes JM, 2007b. The associations between culling due to clinical Johne's disease or the detection of *Mycobacterium avium* subsp. *paratuberculosis* fecal shedding and the diagnosis of clinical or subclinical diseases in two dairy herds in Minnesota, USA. Prev Vet Med. 80: 166–178.

Rankin JD, 1958. The experimental infection of cattle with *Mycobacterium johnei* I. Calves inoculated intravenously. J Comp Pathol. 68: 331–337.

Rankin JD, 1961a. The experimental infection of cattle with *Mycobacterium johnei* II. Adult cattle inoculated intravenously. J Comp Pathol. 71: 6–9.

Rankin JD, 1961b. The experimental infection of cattle with *Mycobacterium johnei* III. calves maintained in an infectious environment. J Comp Pathol. 71: 10–15.

Ridge SE, Baker IM, Hannah M, 2005. Effect of compliance with recommended calf-rearing practices on control of bovine Johne's disease. Aust Vet J. 83: 85–90.

Rossiter, C.A., Hutchinson, L.J., Hansen D. and Whitlock R.H, 1999. Johne's disease prevention/ control plan for beef herds. Manual for Veterinarians. Bovine Prac. 33: 1–22.

Scott HM, Sorensen O, Wu JT, Chow EY, Manninen K, VanLeeuwen JA, 2006. Seroprevalence of *Mycobacterium avium* subspecies *paratuberculosis*, *Neospora caninum*, Bovine leukemia virus, and Bovine viral diarrhea virus infection among dairy cattle and herds in Alberta and agroecological risk factors associated with seropositivity. Can Vet J. 47: 981–991.

Seitz SE, Heider LE, Heuston WD, Bech-Nielsen S, Rings DM, Spangler L, 1989. Bovine fetal infection with *Mycobacterium paratuberculosis*. J Am Vet Med Assoc. 194: 1423–1426.

Sergeant ESG, Nielsen SS, Toft N, 2008. Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds. Prev Vet Med. 85: 92–106.

Sigurðardóttir OG, Valheim M, Press CM, 2004. Establishment of *Mycobacterium avium* subsp. *paratuberculosis* infection in the intestine of ruminants. Adv Drug Deliv Rev. 56: 819–834.

Sockett DC, Conrad TA, Thomas CB, Collins MT, 1992. Evaluation of four serological tests for bovine paratuberculosis. J Clin Microbiol. 30: 1134–1139.

Stabel JR, 1997. An improved method for cultivation of *Mycobacterium paratuberculosis* from bovine fecal samples and comparison to three other methods. J Vet Diagn Invest. 9: 375–380.

Stabel JR, 2006. Host responses to *Mycobacterium avium* subsp. *paratuberculosis*: a complex arsenal. Anim Health Res Rev. 7: 61–70.

Stephan R, Bühler K, Corti S, 2002. Incidence of *Mycobacterium avium* subspecies *paratuberculosis* in bulk-tank milk samples from different regions in Switzerland. Vet Rec. 150: 214–215.

Streeter RN, Hoffsis GF, Bech-Nielsen S, Shulaw WP, Rings DM, 1995. Isolation of *Mycobacterium paratuberculosis* from colostrum and milk of subclinically infected cows. Am J Vet Res. 56: 1322-1324.

Sweeney RW, 1996. Transmission of paratuberculosis. Vet Clin North Am Food Anim Pract. 12: 305–312.

Sweeney RW, Whitlock RH, Rosenberger AE, 1992a. *Mycobacterium paratuberculosis* cultured from milk and supramammary lymph nodes of infected asymptomatic cows. J Clin Microbiol. 30: 166–171.

Sweeney RW, Whitlock RH, Rosenberger AE, 1992b. *Mycobacterium paratuberculosis* isolated from fetuses of infected cows not manifesting signs of the disease. Am J Vet Res. 53: 477–480.

Taylor AW, 1953. Experimental Johne's disease in cattle. J Comp Pathol. 63: 355–367.

Taylor TK, Wilks CR, McQueen DS, 1981. Isolation of *Mycobacterium paratuberculosis* from the milk of a cow with Johne's disease. Vet Rec. 109: 532–533.

Thorel MF, Krichevsky M, Lévy-Frébault VV. 1990. Numerical taxonomy of mycobactin-dependent mycobacteria, emended description of *Mycobacterium avium*, and description of *Mycobacterium avium* subsp. *avium* subsp. nov., *Mycobacterium avium* subsp. *paratuberculosis* subsp. nov., and *Mycobacterium avium* subsp. *silvaticum* subsp. nov. Int J Syst Bacteriol. 40: 254–260.

Tiwari A, Vanleeuwen JA, Dohoo IR, Keefe GP, Haddad JP, Scott HM, Whiting T, 2009. Risk factors associated with *Mycobacterium avium* subspecies *paratuberculosis* seropositivity in Canadian dairy cows and herds. Prev Vet Med. 88: 32–41.

Toft N, Nielsen SS, Jørgensen E, 2005. Continuous-data diagnostic tests for paratuberculosis as a multi-stage disease. J Dairy Sci. 88: 3923–3931.

Tønnesen JJ, 1971. Bekæmpelse af kvægtuberkulose i Jylland. Undersøgelser over den systematiske bekæmpelse af tuberkulosen I jyske kvægbesætninger og over problemerne vedrørende kontrollen med de rensede besætninger (In Danish). A/S Carl Fr. Mortensen, Copenhagen, Denmark, 336 pp.

Uzoigwe JC, Khaitsa ML, Gibbs PS, 2007. Epidemiological evidence for *Mycobacterium avium* subspecies *paratuberculosis* as a cause of Crohn's disease. Epidemiol Infect. 135: 1057–1068.

van Roermund HJ, Bakker D, Willemsen PT, de Jong MC, 2007. Horizontal transmission of *Mycobacterium avium* subsp. *paratuberculosis* in cattle in an experimental setting: calves can transmit the infection to other calves. Vet Microbiol. 122: 270–279.

van Weering H, van Schaik G, van der Meulen A, Waal M, Franken P, van Maanen K, 2007. Diagnostic performance of the Pourquier ELISA for detection of antibodies against *Mycobacterium avium* subspecies *paratuberculosis* in individual milk and bulk milk samples of dairy herds. Vet Microbiol. 125: 49–58.

VanLeeuwen JA, Keefe GP, Tremblay R, Power C, Wichtel JJ, 2001. Seroprevalence of infection with *Mycobacterium avium* subspecies *paratuberculosis*, bovine leukemia virus, and bovine viral diarrhea virus in maritime Canada dairy cattle. Can Vet J. 42: 193–198.

Wang C, Turnbull BW, Gröhn YT, Nielsen SS, 2006. Estimating receiver operating characteristic curves with covariates when there is no perfect reference test for diagnosis of Johne's disease. J Dairy Sci. 89: 3038–3046.

Ward MP, Perez AM, 2004. Association between soil type and paratuberculosis in cattle herds. Am J Vet Res. 65: 10–14.

Weber MF, van Schaik G, 2007. Results of the Dutch bulk milk quality assurance programme for paratuberculosis. Proceedings of the 9th International Colloquium on Paratuberculosis, Oct. 28 – Nov. 2, 2007, Tsukuba, Japan, 324–325.

Weber MF, 2008. Milk quality assurance for paratuberculosis in the national Dutch dairy herd. The Paratuberculosis Newsletter, Sept. 2008, p. 6.

Weber MF, Nielen M, Velthuis AG, van Roermund HJ, 2008a. Milk quality assurance for paratuberculosis: simulation of within-herd infection dynamics and economics. Vet Res. 39: 12.

Weber MF; Verhoeff J, van Schaik G, van Maanen C, 2008b. Surveillance for clinical paratuberculosis: Evaluation of Ziehl-Neelsen test and ELISA through Bayesian modelling. Proceedings of Society for Veterinary Epidemiology and Preventive Medicine, 26–28 March 2008, Liverpool, United Kingdom, p. 262–272.

Wells SJ, Wagner BA, 2000. Herd–level risk factors for infection with *Mycobacterium paratuberculosis* in US dairies and association between familiarity of the herd manager with the disease or prior diagnosis of the disease in that herd and use of preventive measures. J Am Vet Med Assoc. 216: 1450–1457.

Wells SJ, Whitlock RH, Wagner BA, Collins J, Garry F, Hirst H, Lawrence J, Saville WJ, Naugle AL, 2002. Sensitivity of test strategies used in the Voluntary Johne's Disease Herd Status Program for detection of *Mycobacterium paratuberculosis* infection in dairy cattle herds. J Am Vet Med Assoc. 220: 1053–1057.

Wells SJ, Godden SM, Lindeman CJ, Collins JE, 2003. Evaluation of bacteriologic culture of individual and pooled fecal samples for detection of *Mycobacterium paratuberculosis* in dairy cattle herds. J Am Vet Med Assoc. 223: 1022–1025.

Whitlock RH, Rosenberger AE, Siebert M, Sweeney R, 1992. Environmental contamination on dairy farms with cattle infected with *Mycobacterium paratuberculosis*. Proc 17th World Buiatrics Congress, St Paul, Minnesota, USA, vol. 1, p. 84–88.

Whitlock RH, Rosenberger AE, Sweeney RW, Spencer PA, 1996. Distribution of *M. paratuberculosis* in tissues of cattle from herds infected with Johne's disease. Proceedings of the Fifth International Colloquium on Paratuberculosis. September 29 – October 4, 1996, Madison, Wisconsin, USA, p. 168–174.

Whittington RJ, Windsor PA, 2009. In utero infection of cattle with *Mycobacterium avium* subsp. *paratuberculosis*: A critical review and meta-analysis. Vet J. 179: 60–69.

Wilson DJ, Rossiter C, Han HR, Sears PM, 1993. Association of *Mycobacterium paratuberculosis* infection with reduced mastitis, but with decreased milk production and increased cull rate in clinically normal dairy cows. Am J Vet Res. 54: 1851–1857.

Yekutiel P, 1980. Eradication of infectious diseases: A critical study. In: Contribution to Epidemiology and Biostatistics, vol. 2, Karger, Basel, Switzerland, 164 pp.

Yokomizo Y, Yugi H, Merkal RS, 1985. A method for avoiding false-positive reactions in an enzyme-linked immunosorbent assay (ELISA) for the diagnosis of bovine paratuberculosis. Jpn J Vet Sci. 47: 111–119.

Accompanying papers

Diagnosis

I) Nielsen SS, Toft N, 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques. Vet Microbiol. 129: 217–235.

II) Nielsen SS, Ersbøll AK, 2006. Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows. J Dairy Sci. 89: 4557–4566.

III) Toft N, Nielsen SS, Jørgensen E, 2005. Continuous-data diagnostic tests for paratuberculosis as a multi-stage disease. J Dairy Sci. 88: 3923–3931.

IV) Nielsen SS, Toft N. 2006. Age-specific characteristics of ELISA and fecal culture for purpose specific testing for paratuberculosis. J Dairy Sci. 89: 569–579.

V) Nielsen SS, 2008. Transitions in diagnostic tests used for detection of *Mycobacterium avium* subsp. *paratuberculosis* infections in cattle. Vet Microbiol. 132: 274–282.

VI) Nielsen SS, Krogh MA, Enevoldsen C, 2009. Time to occurrence of drop in milk production in cows with various paratuberculosis antibody profiles. J Dairy Sci. 92: 149–155.

Prevalence

VII) Nielsen SS, Toft N, 2009. A review of prevalences of paratuberculosis in farmed animals in Europe. Prev Vet Med. 88: 1–19.

VIII) Sergeant ESG, Nielsen SS, Toft N, 2008. Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds. Prev Vet Med. 85: 92–106.

IX) Nielsen SS, Toft N, Jørgensen E, Bibby BM, 2007. Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response. Prev Vet Med. 81: 290–305.

Risk factors

X) Nielsen SS, Gröhn YT, Quaas RL, Agger JF, 2002. Paratuberculosis in dairy cattle: Variation of the antibody-response in offspring attributable to the dam. J Dairy Sci. 85: 406–412.

XI) Nielsen SS, Bjerre H, Toft N, 2008. Colostrum and milk as risk factors for infection with *Mycobacterium avium* subsp. *paratuberculosis* in dairy cattle. J Dairy Sci. 91: 4610–4615.

XII) Nielsen SS, Toft N, 2007. Assessment of management-related risk factors for paratuberculosis in Danish dairy herds using Bayesian mixture models. Prev Vet Med. 81: 306–317.

Control

XIII) Kudahl AB, Nielsen SS, Østergaard S, 2008. Economy, efficacy and feasibility risk-based control program against paratuberculosis. J Dairy Sci. 91: 4599–4609.

Paper I

Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques

Nielsen SS, Toft N

Published in
Veterinary Microbiology 2008, 129, 217–235

Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques

S. S. Nielsen[1], N. Toft[2]

[1]*Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark*
[2]*Danish Meat Association, Vinkelvej 11, DK-8620 Kjellerup, Denmark*

ABSTRACT

Infections with *Mycobacterium avium* subsp. *paratuberculosis* (MAP) can be latent for years without affecting the animal, but the animal may become infectious or clinical at some point. Diagnosis of paratuberculosis can be a challenge primarily in latent stages of the infection, and different diagnosis interpretations are usually required by the variety of decision makers. The objective of this paper was to provide a critical review of reported accuracies of ELISA tests, interferon-γ assays (IFN-γ) and faecal culture (FC) techniques used for diagnosis of three defined target conditions: MAP infected, MAP infectious and MAP affected animals.

For each animal species, target condition and diagnostic test-type, sensitivities (Se) and specificities (Sp) were summarised based on a systematic, critical review of information in literature databases. The diagnostic test information often varied substantially for tests of the same type and make, particularly ELISA, which was the most frequently reported test-type. Comparison of the various tests accuracies was generally not possible, but stratification of test-evaluations by target condition improved the interpretation of the test accuracies. Infectious and affected animals can often be detected, but Se for infected animals is generally low. A main conclusion of the review was that the quality of design, implementation and reporting of evaluations of tests for paratuberculosis is generally poor. Particularly, there is a need for better correspondence between the study population and target population, i.e. the subjects chosen for test evaluation should reflect the distribution of animals in the population where the test is intended to be used.

1. INTRODUCTION

Paratuberculosis is a chronic infection, which has been of particular concern in ruminants. The infection is caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). The major effects of the infection on the animal level can be reduced milk yield (Benedictus et al., 1987; Kudahl et al., 2004), premature culling and reduced slaughter value (Benedictus et al., 1987), and losses due to continued spread of the infection (Kudahl et al., 2007). Not all infected animals will experience losses, which may be because of culling for other reasons than paratuberculosis or because they can resist the infection developing into the debilitating stages. The occurrence of the latter is still poorly understood (Mortensen et al., 2004).

Prevalences of the infection vary world-wide (Kennedy and Benedictus, 2001), but most notably the apparent prevalences vary by the test and test strategies used in the prevalence studies conducted. Control of the infection can be obtained via timely detection and culling of infectious animals and reduction of transmission from these animals. Eradication will usually require the detection and isolation of infected animals, as these potentially can become infectious at some point in time. “Isolation” in this regard means that infected animals and their excretions should not be allowed contact with susceptible animals. Eradication is

defined as: “The purposeful reduction of specific disease prevalence to the point of continued absence of transmission within a specified area by means of a time-limited campaign” (Yekutiel, 1980). This is to emphasize that complete eradication would require eradication of a microbial agent globally. Control is described as “any effort directed toward reducing the frequency of existing disease to levels biologically and/or economically justifiable or otherwise of little consequence” (Martin et al., 1987).

The objective of this report was to conduct a critical review of reported diagnostic test evaluations of ELISA, FC and interferon-γ tests (IFN-γ) used for ante mortem diagnosis of conditions caused by infection with MAP. To facilitate the comparison across test studies, three target conditions: affected, infectious and infected with MAP were defined and diagnostic sensitivity (Se, probability of correct test positive classification) and diagnostic specificity (Sp, probability of correct test negative classification) reported with respect to these conditions.

2. STAGE OF INFECTION / CONDTIONS DETECTED

2.1. Pathogenesis

Diagnosis and thereby control of the infection are hampered by a long incubation period. It is generally assumed that infections with MAP occur in young animals, and that some age-resistance occurs. Cattle are thought to be most susceptible from 0 to 4 months of age (Taylor, 1953), although infections have been established in adults fed high dosages of MAP (Doyle, 1953). Similar conditions can be speculated to occur for other animal species. Clinical disease has been observed to most frequently occur among cattle 2–5 years of age, although cattle from very young to very old (0–13 years of age) have been affected (Doyle and Spears, 1951).

MAP is an intracellular pathogen. Subsequent to infection, MAP is initially controlled by a predominating T helper 1 (Th1) response. Th1-cells are, among other features, characterised by their production of interferon-γ and some IgG2. Later in the course of infection, a predominant Th2 response may occur, and control of the infection is thought to be lost (Stabel, 2000). During the Th1 response, Map is shed in small numbers, which may be sufficient to elicit a positive result in faecal culture tests (FC). There is correlation between occurrence of IgG and bacterial shedding of MAP (Nielsen and Toft, 2006a), but the time-wise relations between the two events is not fully described. Experimental infections suggest that bacterial shedding decreases 10–14 months after inoculation, to increase again later, with sero-conversion occurring around 10 months post-inoculation (Lepper et al., 1989). Waters et al. (2003) demonstrated both cellular and humoral immune responses approximately 100–150 days after initial infections with MAP, and Eda et al. (2006) also demonstrated IgG in calves less than 1 year after inoculation. In naturally infected animals, sero-conversion has been shown to occur in 95–98% of cows shedding MAP (Nielsen and Ersbøll, 2006). The age at which sero-conversion occurred was from 2.2 to 11.7 years. It is speculated that great variation of the time to occurrence of bacterial shedding and occurrence of antibody responses are caused by variation in infective doses occurring with natural MAP infections. Studies using fixed dosages and known times of infection have also resulted in great variation in the time to occurrence of FC-positivity and ELISA-positivity (Lepper et al., 1989), and the temporal variation in pathogenetic events may be further enhanced if the size and number of dosages varies. In test-evaluations, it is therefore needed to consider the stage of infection. Generally, age can be an indicator of the stage of

infection, in that young animals will rarely be expected to shed detectable amounts of bacteria and have IgG1, whereas older animals are more likely to have bacterial shedding, antibodies and clinical disease. The age-distribution among study objects can therefore be of great significance in the evaluation of a diagnostic test. As an example, the probability of detecting infected cows 2 years of age using an antibody ELISA has been estimated to 0.06, whereas the same probability was 0.50 for cows 5 years of age (Nielsen and Toft, 2006a). The distribution of animals among different stages of the infection can also be of great importance.

Different antigens are used in various immunological tests. The antigens are usually derived from MAP or *M. avium* subsp. *avium* (MAA). In the paratuberculosis literature, MAP strain 18 is MAA serovar 2 (Chiodini, 1993). Different preservations of the antigens are being used based on their biological characteristics (Koets et al., 2001) and availability. Generally, antigens should be immunogenic to be used for diagnostic tests. Cho and Collins (2006) showed that proteins derived from culture filtrates rather than cellular extracts were more likely to be antigenic. The use of antigens from MAA may be as useful as antigens from MAP (Nielsen et al., 2001), but there have been made no extensive comparisons of antigens on field data indicating which antigens are the better. Irrespective of which of the two bacterial species the antigen is derived from, both immunogenecity and cross-reactivity should be considered. So far, the superiority of one specific antigen has not been demonstrated, where both immunogenecity and cross-reactivity have been evaluated. It is likely that this may be specific to specific geographic areas, due to potentially different distribution of bacteria potentially causing cross-reactions, such as MAA, from one area to the other. In herds with high sero-prevalences estimated by two commercial ELISAs (Parachek, CSL/Biocor, Omaha, NE, USA and HerdChek, IDEXX Laboratories Inc., Westbrook, ME) the prevalence of environmental mycobacteria were higher compared to herds with a low sero-prevalence, suggesting that a high prevalence of environmental mycobacteria may result in many false-positive ELISA results (Roussel et al., 2007). It has also been demonstrated that experimental infections with environmental mycobacteria such as *M. intracellulare*, *M. scrofalceum* and *M. terrae* can result in significant serological reactions (Osterstock et al., 2007).

2.2. Target conditions

The target condition detected by any diagnostic test is essential in the evaluation of the diagnostic accuracy of the test. Some authors refer to this a “case definition” (e.g. Collins et al., 1991). The choice of target condition for the diagnosis “paratuberculosis” should vary depending on the purpose of testing, i.e. the effects that are of primary interest to decision makers. A schematic presentation of the pathogenesis and the effects is given in Fig. 1. From this, a number of conditions can be identified. The target condition chosen by the evaluators in a given study ideally depends on the decision makers, whom can subsequently make decisions based on the test results. In this report, three target conditions, affected, infectious and infected with MAP, have been defined to classify the test evaluation studies included in the review. These target conditions are considered pivotal.

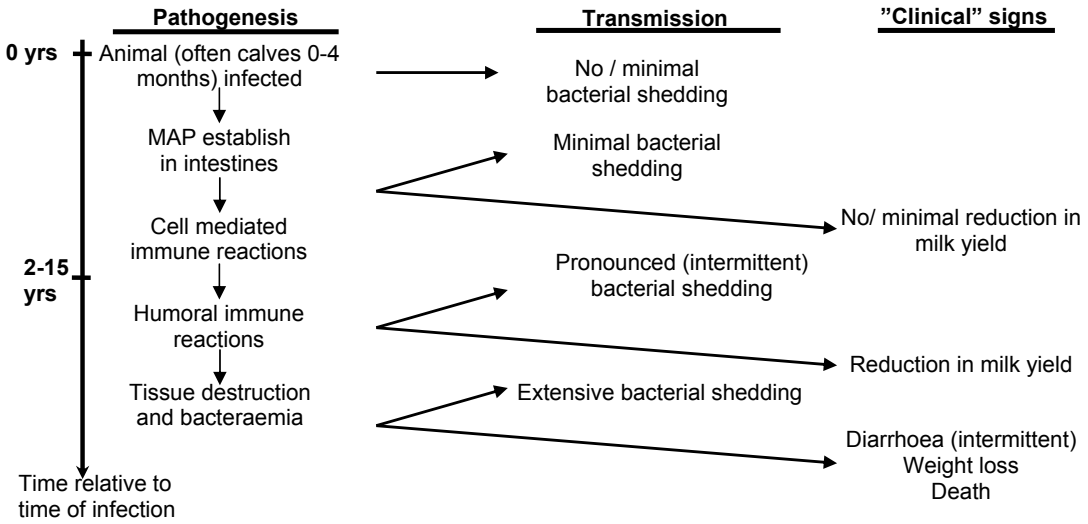


Fig. 1. Schematic presentation of various stages of infection and their effects. This presentation may represent the typical picture, but deviations are likely to occur

2.2.1. Animals affected by MAP.

Animals affected by MAP are usually classified based on clinical signs such as diarrhoea (persistent or intermittent), chronic weight loss or reduced milk production. MAP infection should be present, which could be documented via gross pathology, histopathology or cultivation of MAP from tissues or faeces. The animal does not have control of the infection and is affected to a degree, where parameters like milk production and general performance is decreasing due to the infection.

This definition can be of use for a decision maker whom wishes to make decisions based on the performance of the cow. Reduction in milk yield is often not recorded and reported systematically in the studies, and this aspect of the infection is therefore rarely included.

2.2.2. Animals infectious with MAP

Infectious animals are defined as those that shed MAP at the time of testing with the test under evaluation and thereby they are a risk for transmission of MAP to susceptible herd-mates. The condition "Infectious" also includes animals which are "Affected". In principle, the infectious group may also contain non-infected animals, which may be passive shedders of MAP. These animals are hypothesised to ingest MAP from heavy environmental contamination (Sweeney et al., 1992) without being infected. The "infectious" status of these animals would therefore be considered to be transient.

In many study reports, only shedding in faeces is included. The shedding is defined based on one or more tests evaluating the presence of MAP in faeces. Animals which are transmitting the agent in milk and in uteri are thereby not included specifically. Given that animals are infectious via milk and in uteri without shedding bacteria in faeces, they will therefore bias the accuracy estimated in the studies reported.

2.2.3. Animals infected with MAP.

Infected animals carry MAP intracellularly but substantial replication need not take place because the infection can be latent. The condition “Infection” also includes animals which are “Infectious” and “Affected”. The definition of “infection with MAP” is any condition where entrance and persistence of MAP have lasted long enough to give an immune response at any time during their life; i.e. there is no time-specific cut-off for this event to occur. It is assumed that once a cow has an established infection, the infection persists for life.

2.2.4. Animals not affected, infectious or infected with MAP

The diagnostic Se of a test reflects the ability to detect the target condition given it is present and the diagnostic Sp of a test reflects the ability to test negative given the target condition is not present. In many studies, the target conditions mentioned above are mixed. Studies defining one target condition for the evaluation of Se and another for Sp can therefore be subject to peculiar interpretations. The interpretation of “false-positives” can therefore vary.

The Sp of the condition “non-affected” has to our knowledge not been systematically assessed in any studies.

The Sp of the condition “non-infectious” refers to an animal, which does not shed MAP on the time of testing, but some of these non-infectious animals can be infected. Therefore, false-positive test-results include both infected as well as non-infected (i.e. MAP-free) animals. The decision maker will usually have little use of this information, unless documentation exists to show near perfect Sp of the test for non-infected animals. This would mean that all false-positive animals are indeed infected, but not infectious at the time of testing.

The Sp of the condition “non-infected” is the situation where an animal is free of the infection. False-positive reactions are due to cross-reactions to other mycobacteria, laboratory errors, vaccination reactions and the like for immunological tests (Houe et al., 2004). False-positive test results could also be a consequence of the “pass-through” phenomenon, where ingested bacteria are shed 1–7 days post-ingestion, potentially without being infected (Sweeney et al., 1992).

2.2.5. Utility of test-results related to different conditions

The utility of the test results related to the three conditions defined above could be as follows:

- Affected animals. The value of these animals is low because their production is reduced, they loose weight and the value at slaughter is probably reduced. There is a risk they will die from the infection.
- Infectious. These animals are currently infectious and are a risk to susceptible animals. They should be managed so that transmission to susceptible animals is avoided. These animals constitute both a short-term and a long-term economic burden, as they are likely to have a reduced milk yield or will experience it in the near future. The long-term loss will be due to their transmission of MAP to herd-mates.
- Infection. Infected animals constitute a risk of becoming infectious and thereby transmit the infection to susceptible animals. A population containing infected animals cannot be declared free of infection and proper identification of infected animals is important in herd-certification schemes, when trying to establish ‘MAP-free’ herds/populations or keeping MAP out of certified herds. In economic analyses, these

animals may be of interest in long-term planning only, as their effect on the population will often be seen only after a number of years.

Theoretically, the conditions may be easily defined. In practice, this may not be the case. Other conditions can be defined, depending of their use, e.g. assuming that the cell-mediated immune reactions are studied, it is necessary to establish a condition “occurrence of cell-mediated immune reactions”.

2.2.6. Effect of age on the condition detected.

Considering the chronic nature of the infection, with many disease stages, it could be speculated that older animals are of greater risk of having a given condition than younger animals. However, this does not necessarily imply that it is easier to detect the condition in older animals. Higher diagnostic Se for detection of infection has been demonstrated for older cows relative to younger cows (Nielsen and Toft, 2006a), but age does not seem to influence detection of infectious and affected cattle (Nielsen and Toft, 2006a; 2006b]), although this has not been formally tested.

3. THE IDEAL TEST EVALUTION

Very few if any perfect test evaluations have yet been performed. The lack of a 100% accurate reference test and a variable incubation time seem to be primary obstacles in doing so. The effect of choice of reference test on accuracy estimates was demonstrated by McKenna et al. (2005), where accuracy estimates of ELISA based on use of tissue culture as reference test was different from accuracy estimates based on faecal culture as reference test. “Infection” may be established if a thorough microbiological examination of the animal is performed at slaughter, but it is insufficient to sample tissues only from ileum and the ileocecal lymph nodes, because this sampling procedure will fail in detecting of many infected animals (Whitlock et al., 1996). The authors state that up to 100 sites per animal are required sampled to establish the true infection status of the animal. A study containing an adequate number of animals would therefore become quite expensive, since the prior infection status should not be known when carrying out the study, in order to ascertain a distribution of infection stages that is representative of the distribution in the target population.

Greiner and Gardner (2000) provide an extensive check list, which should be used as a starting point for any epidemiological validation of diagnostic tests. From their recommendations and the discussion of target conditions in the previous sections, we suggest that any test evaluation must at least include the following components:

(a) Data must be from an observational study and the study population should be representative of the target population in which the test is to be used, i.e. the variations in incubation period, exposure dosages and age-distribution of the target population should be reflected in the study population.

(b) The target condition should reflect the intended purpose of the test. The same target condition should be used for evaluation of both the Se and Sp. The interpretation of the test-information should be made in concordance with the target condition, and the interpretation of false-positives needs to be clear. As an example, the target condition “Infected” can be studied by classical test-evaluation methods, where animals are classified into truly infected, not infected based on microbiological examination of multiple tissues per animal. An

alternative is to use latent class analyses (e.g. Nielsen et al., 2002), where tests that are biologically unrelated are studied, e.g. FC for detection of MAP and antibody ELISA for detection of humoral immune reactions.

(c) A calculated sample size which reflects the purpose of the test evaluation, i.e. the choice of precision in the estimates should depend on the specific purpose of the test evaluation.

4. FREQUENT MISTAKES IN TEST EVALUATIONS PERFORMED

A number of mistakes are usually made in evaluation of diagnostic tests. Some of these can be avoided by thorough planning of the study, whereas others are difficult to avoid while still making the study practically and economically feasible. Frequent mistakes include:

(1) Selection of animals for the evaluation is performed by use of the test being evaluated, or a test measuring the same response. Examples include animals selected from herds classified free of infection based on negative humoral immune responses, where the test evaluated detects the same responses.

(2) Variable case definitions used to classify animals across a study. All study objects should be subject to the same classification procedures. An example is the study by Sockett et al. (1992b), in which not all infected animals were subject to the same set of classification procedures, since only sero-positive and FC-negative animals were subject to the defining histopathological procedures. This would result in an over-estimation of the Se, because animals in early-stages of infection could be missed by the reference test. However, the study could be included for inference on infectious animals, given that FC was used to define the reference status, because all animals had been subject to the FC.

(3) Another frequent, but to some extent unavoidable mistake is the use of a study population from a region historically known to be free of MAP to estimate Sp and sample the non-free target population to estimate Se. These two populations should be geographically comparable, because the environmental flora can be expected to give rise to different cross-reactive responses in serological tests.

5. REVIEW OF TEST EVALUTIONS IN LITERATURE

A review of test-evaluations was carried out by searching the available databases by 14 November 2006. These included: Agricola (1970 to September 2006); Agris (1975 to September 2006); Biological Abstracts/RPM (R) (1989–2003); BIOSIS Previews (2004 to 11 November 2006); Biological Abstracts (1990–2000); CAB Abstracts (1973 to September 2006) and Medline through Pubmed (1970 to 14 November 2006). The search terms were: paratuberculosis, Johne's or Johnes combined with diagnosis, diagnostic accuracy, Se, Sp, validation or diagnostic performance.

This search generated 2137 hits including duplicate records. After exclusion of duplicate records and non-peer-reviewed publications, 312 publications remained. These were further reduced to 102 publications by exclusion of studies, where the abstract indicated that a test-evaluation was not an objective of the publication or if the language was not English. In cases of uncertainty on whether a test-evaluation had been performed, the studies were included for further assessment of the full paper. The remaining 102 publications were further evaluated for study objective and data quality.

5.1. Data extraction

Data were extracted from the publications in a standard form including the following items:

a) Animal species;
b) Type of test, i.e. serum antibody ELISA, milk antibody ELISA, FC and IFN-γ;
c) Test system, i.e. name of test for commercial tests and in-house for non-commercial tests;
d) Antigen used in serological tests;
e) Study design: Observational (case-control; cohort; cross-sectional) or experimental; longitudinal or instantaneous;
f) Data origin: field data or serum-bank or similar;
g) Conditions detected: “infectious” if animals were shedding bacteria; “infected” if animals were deemed infected by a reference method stated; “affected” if cows were clinical, had reduced production or were in some other way deemed to be affected by the infection (“shedding” alone was not considered affected); “from free region” if this had been justified (herds that had once been FC negative were not from “a free region”); “non-infectious” were from non-shedding animals);
h) Sample size, i.e. the number of animals assessed among animals with condition (for Se) and without condition (for Sp);
i) The number of test-positives among animals with condition and the number of test-negatives among animals without the condition;
j) Uncertainty measures, e.g. 95% confidence interval (95% CI) for classical methods or 95% credibility posterior intervals for Bayesian methods;
k) Cut-off used for discriminating between positive and negative animals in tests that are not dichotomous;
l) Age distribution of the study population(s);
m) Arguments for inclusion and exclusion of the data from the study;
n) Authors and year of publication;
o) Year of study;
p) Geographical origin of data.

5.2. Assessment of study quality

Many studies included some sort of bias, and it was not possible to avoid these biases generally. Therefore, it was decided to include studies if the following criteria were fulfilled:

a) One of the objectives of the study should be related to test evaluation;
b) A unique target condition should be defined for animals with the condition and animals without the condition. The target condition could differ between the estimates for Se and the estimates for Sp. If a condition could not be determined or varied across study objects, the study was excluded;
c) Animals were classified by the same criteria within the study;
d) The study population was not selected or defined by use of the test (or a related test) under evaluation, e.g. selection of herds that had previously been shown to be negative by serological methods could not be used for evaluation of a serological test;
e) Random inclusion of study objects was used, e.g. studies where study objects were included due to being suspect were excluded from the current study. Given that the selection of study objects was not described sufficiently, these studies were also excluded;

f) Infections were natural, i.e. results from experimental infections were excluded, but results of natural infections were included if these were given separately in a given publication;
g) The test-result of a given test under evaluation was based on a single test, i.e. repeated test-results could be used for defining a given condition but could not be used for the test evaluated;

Some test-results were confirmed by re-testing, but the original data were used. The argument for doing so is that by sufficient re-testing, the required result can usually be obtained if enough re-tests are carried out, and the same set of procedures should be applied to all samples.

5.3 Data analysis

The recorded data from each of the studies were tabulated in separate tables for different animal species. The target condition detected, age and test-characteristics were included in the tables, and the range of estimates was extracted.

6. TEST EVALUATION SUMMARIES

The 102 publications evaluated in detail contained a total of 153 test-evaluations. A test-evaluation is here defined as the evaluation of one test, but some publications contained more than one test. In some publications, more than one test or more than one condition was evaluated. In Table 1, the distribution of 153 test-evaluations is given by target condition and animal species. A total of 68 studies were excluded, as they did not fulfil the inclusion criteria. The remaining studies included 58 studies from cattle, 15 studies from goats, 11 studies from sheep and 1 study from deer. Four of these studies were conducted using latent class analyses: 1 on cattle, 2 on goats and 1 on sheep. The distribution among tests used was a follows: 64 serum antibody ELISA (SELISA) studies, 6 milk antibody ELISA (MELISA) studies, 4 IFN-γ studies and 8 studies on FC. An overview of these is given in the subsequent sections, with division into animal species, test-type and condition detected. The test-evaluations did not cover all types of tests FC, SELISA, MELISA and IFN-γ for all animal species, and studies on those not mentioned have not been reported or did not fulfil the inclusion criteria set.

Table 1. Overview of 102 publications on evaluation of diagnostic test for paratuberculosis in animals, divided into three target conditions and animal species

Condition	Cattle	Goats	Sheep	Llamas & alpacas	Deer	Water buffaloes	Total
Affected	4	8	5	0	1	0	18
Infectious	36	2	0	0	0	0	38
Infected	18	5	6	0	0	0	29
Excluded	45	9	12	1	0	1	68
Total	103	24	23	1	1	1	153

Although data should have been extracted as specified in Section 5.1, some information was consistently not available from most publications. Therefore, only parameters shown in Table 2 and Table 3 are reported further, as these were almost consistently reported. Age of the study groups was also included, although this information was only reported in some cases.

Unless otherwise stated, it is assumed that adult, mature animals comprised the target population.

6.1. Cattle – Faecal Culture

Usually, the Sp of FC is considered to be almost 100%, if the isolates obtained at culture are confirmed to be MAP by molecular methods such as confirmatory IS900 PCR. However, due to the potential pass-through phenomenon (Sweeney et al., 1992), it is possible that testing of non-infected animals on contaminated premises can lead to false-positive reactions. A latent-class approach to evaluation of the Sp of FC for detection of non-infected animals would therefore seem appropriate. Nielsen et al. (2002) estimated the Sp to be 98% in a population, where the non-infected animals were subject to contamination from infected herd-mates.

The Se from the test-evaluations in the present study are given in Table 2, for the three different target conditions. The Se of FC to detect affected animals have been estimated in one study only, resulting in an estimated of 0.70, which is similar to the Se of 0.74 estimated for detection of infectious animals. The Se's of FC for detection of infected animals were in the range 0.23–0.29, except for the study by Billman-Jacobe et al. (1992). However, in that study the infection status was determined at necropsy of cull cattle, many of which where clinically ill or gave strong positive reactions in serological tests. Therefore, their estimate must be considered biased to a degree which suggests it should be excluded from the review or that the true target condition of the study differ from their reported target condition.

Table 2. Sensitivity of faecal culture methods used for diagnosis of affected, infectious and infected cattle.

Condition	Test medium[#]	No. with condition	No. test-Positive	Sensitivity (95% C.I.)	Age	Reference
Affected	HEYM	56	39	0.70 (0.56, 0.81)	?	Egan et al., 1999
Infectious	HEYM	111	82	0.74 (0.65, 0.82)	> 2 yrs	Sockett et al., 1992b
Infected	HEYM	177	86	0.49 (0.41, 0.56)	?	Billman-Jacobe et al., 1992
Infected	TREK	160	36	0.23 (0.16, 0.30)	?	McKenna et al, 2005
Infected	HEYM	321	79	0.25 (0.20, 0.30)	All cattle in herd, >0 yr	Whitlock et al., 2000
Infected	HEYM	232	67	0.29 (0.23, 0.35)	All parturient cows	Whitlock et al., 2000

[#]) HEYM = Herrold's Egg Yolk Medium; TREK = TREK ESP culture system, TREK Diagnostics, Cleveland, Ohio, USA

6.2. Cattle - Serum antibody ELISA

The serum antibody ELISA for cattle is the test most frequently evaluated. The test evaluations include studies on a number of commercial ELISAs and in-house ELISAs, a variety of antigen preparations and different age-groups of animals. In Table 3, the Se and Sp is given for each study. The test used is classified into groups based on the producer and the antigen used. The antigen preparations are in most cases not comparable. Therefore, the names in the groups indicate only the source of the antigen. It is assumed that the antigen in the Parachek® (Prionics AG, Schlieren-Zurich, Switzerland) test is from the MAP VRI 316 strain, but this has not been confirmed. The producer did not respond to requests of the test specification, but the test should apparently be based on the test described originally by Milner et al. (1990).

Table 3. Sensitivity and specificity estimates from studies of serum antibody ELISAs for detection of affected (A), infectious (I) and infected (E) cattle

Condition		Test	Antigen	Sample size[#]		Test outcome[¤]		Se[§]	Sp[§]	Age	Reference
Se[§]	Sp[§]			C+	C-	T+|C+	T-|C-				
A	NA	Various	ATCC 19698	40		20		0.50		≥12 mo.	Bech-Nielsen et al., 1992
A	NA	HerdChek	IDEXX	62		54		0.87		?	Sweeney et al., 1995
A	NA	Parachek	VRI316	56		43		0.77		?	Egan et al., 1999
I	I	Various	ATCC 19698	134	62	64	62	0.48	1.00	≥12 mo.	Bech-Nielsen et al., 1992
I	I	HerdChek	IDEXX	373	2383					Cows	Berghaus et al., 2006
I	I	HerdChek	IDEXX	72	617	53	540	0.74	0.88	?	Hendrick et al., 2005
I	I	HerdChek	IDEXX	44	607	11	569	0.25	0.94	?	Stabel et al., 2002
I	I	HerdChek	IDEXX	174		62		0.36		?	Sweeney et al., 1995
I	I	HerdChek	IDEXX	41	263	13	258	0.32	0.98	?	Sweeney et al., 1995
I	I	Various	LAM	102	65	61	54	0.60	0.83	?	Sweeney et al., 1994
I	I	Various	PPA3	67	513	42	460	0.63	0.90	≥ 2 yr	Klausen et al., 2003
I	I	Various	PPA3	8	16	5	11	0.63	0.69	?	Paolicchi et al., 2003
I	I	Various	PPA3	177	196	71	187	0.40	0.95	?	Sockett et al., 1992a
I	I	Various	Various	60	44	37	18	0.62	0.41	?	Abbas et al., 1983
I	I	Various	Various	60	44	50	39	0.83	0.89	?	Abbas et al., 1983
I	I	Various	Various	36	156	34	129	0.94	0.83	?	Colgrove et al., 1989
I	I	Various	Various	14	210	10	174	0.71	0.83	> 6 mo	Spangler et al., 1992
I	I	Parachek	VRI316	170	1751	40	1719	0.24	0.98	≥ 2nd lact.	Lombard et al., 2006
I	I	Parachek	VRI316	177	196	61	194	0.34	0.99	?	Sockett et al., 1992a
I	I	Various	VRI316	60	304	22	287	0.37	0.94	?	Eamens et al., 2000
I	E	HerdChek	IDEXX	415	359	127	342	0.31	0.95	?	Collins et al., 2005
I	E	HerdChek	IDEXX	198	346	48	345	0.24	1.00	>2 yr	Kalis et al., 2002
I	E	SVANOVA	LAM	15	100	6	91	0.40	0.91	2 to 15 yr	Glanemann et al., 2004
I	E	Various	PPA3	64	68	51	61	0.80	0.90	?	Nielsen et al., 2001
I	E	IDEXX Scand	Various	198	346	66	322	0.33	0.93	>2 yr	Kalis et al., 2002
I	E	IDEXX Scand	Various	198	346	51	335	0.26	0.97	>2 yr	Kalis et al., 2002
I	E	Pourquier	Various	415	359	116	359	0.28	1.00	?	Collins et al., 2005
I	E	SERELISA	Various	301	359	134	304	0.45	0.85	?	Collins et al., 2005
I	E	Various	Various	156	200	126	200	0.81	1.00	>15 mo.	Yokomizo et al., 1991
I	E	Parachek	VRI316	126	196	71	194	0.56	0.99	?	Collins et al., 1991
I	E	Parachek	VRI316	415	359	118	358	0.28	1.00	?	Collins et al., 2005
I	E	Parachek	VRI316	64	68	51	61	0.80	0.90	?	Nielsen et al., 2001
I	NA	Various	Various	106	341	50	340	0.47	1.00	?	Reichel et al. 1999
I	NA	Parachek	VRI316	106	341	33	334	0.31	0.98	?	Reichel et al. 1999
E	E	HerdChek	IDEXX	160	834	14	814	0.09	0.98	?	McKenna et al, 2005
E	E	Various	LAM	22	378	4	363	0.18	0.96	?	McNab et al., 1991
E	E	SVANOVA	Various	160	834	27	757	0.17	0.91	?	McKenna et al, 2005
E	E	Parachek	VRI316	160	834	11	801	0.07	0.96	?	McKenna et al, 2005
E	NA	Parachek	VRI316	1188		265		0.22		>2 yrs	Jubb et al., 2004
NA	E	Parachek	VRI316		15566		15467		0.99	?	Holmes et al., 2004
NA	E	Parachek	VRI316		5588		5579		1.00	?	Jubb and Galvin, 2004
NA	E	Parachek	VRI316		1049		1028		0.98	Cows	Pitt et al., 2002

[§]) Se= Sensitivity; Sp=Specificity; A=Affected; I=Infectious; E=Infected; NA=Not available.

[#]) Sample sizes for animals with condition (C+) and without condition (C-)

[¤]) No. of test-positive (T+) or test-negative (T-) given the condition (C)

IDEXX Laboratories, Inc. (Westbrook, ME, USA) being the provider of the IDEXX HerdChek *Mycobacterium paratuberculosis* test, cannot share the specification of their antigen preparation either, and it is unknown whether the both the HerdChek and the Parachek test have remained the same over the years. Other test-names were categorised as “Various” because they often do not have a name. For antigen preparations, the variety of preparations gives reason to the same “Various” group. The test names and antigen preparations are therefore not suitable for further subdivision of the data. In most studies, the age-groups studied are incompletely characterised. For those studies in which they were given, it is clear that the studies are hardly comparable, but it is assumed that the majority of animals in each study are a random collection of parturient animals, except if stated otherwise.

6.3. Cattle - Milk antibody ELISA

The constitution of milk antibody ELISAs has been less variable than the serum antibody ELISA. Only two different antigens have been used: a lipo-arabinomannan preparation (LAM) and a commercially available antigen from Allied Monitor (Fayette, MO, USA). The Se and Sp still vary, which may be primarily due to the choice of cut-off used in the different studies. A summary is given in Table 4. The test has not been evaluated for affected animals and very few studies have been conducted on infected animals. Therefore, the most studies have been conducted for diagnosis of infectious animals, and the Se varies from 0.29 to 0.61, with Sp in the range of 0.83–1.00.

Table 4. Sensitivity and specificity estimates from studies of milk antibody ELISAs for detection of infectious (I) and infected (E) cattle

Condition		Test	Antigen	Sample size[#]		Test outcome[¤]		Se[$]	Sp[$]	Age	Reference
Se[$]	Sp[$]			C+	C-	T+|C+	T-|C-				
I	I	Antel[$]	Allied	72	617	44	584	0.61	0.95	?	Hendrick et al., 2005
I	I	Various	Allied	67	513	36	487	0.54	0.95	≥ 2 yr	Klausen et al., 2003
I	I	Antel[$]	Allied	170	1751	36	1724	0.21	0.98	≥ 2nd lact	Lombard et al., 2006
I	I	Various	LAM	102	65	61	54	0.60	0.83	?	Sweeney et al., 1994
I	E	Antel[$]	Allied	364	352	105	351	0.29	1.00	?	Collins et al., 2005
E	E	Various	Allied	2662 comb.				0.39	0.96	Cows	Nielsen et al., 2002

[$]) Se= Sensitivity; Sp=Specificity; I=Infectious; E=Infected.

[#]) Sample sizes for animals with condition (C+) and without condition (C-)

[¤]) No. of test-positive (T+) or test-negative (T-) given the condition (C)

[$]) Antel Biosystems Inc., Lansing, Michigan, USA

6.4. Cattle – interferon-γ

The IFN-γ has only been evaluated in two studies in cattle ([Paolicchi et al., 2003] and [Huda et al., 2004]), and on a limited data material. The results are summarised in Table 5, and the Se for detection of infectious animals varied from 0.13 to 0.85. Unfortunately, the results may not be representative of the populations in general, but so far the studies are the only information available. In Paolicchi et al. (2003), the data material is based on only one herd, in which clinical paratuberculosis had occurred. In Huda et al. (2004), the negative reference herds were chosen among herds with low prevalence of antibody positive cows. It is uncertain whether both publications should be excluded based on these weaknesses. An advantage of the study by Huda et al. (2004) is the separation into three age-groups (Table 5). Neither of the studies assessed the Se of the test for detection of infected animals.

6.5. Goats – Faecal Culture

Kostoulas et al. (2006) have conducted the only study reporting on the accuracy of FC in goats. The condition studied was infected animals, and Se and Sp were assessed using latent class methods on data derived from animals >1 year of age. The Se was estimated to 0.08 (95% credibility posterior interval (95% CPI): 0.02; 0.17) and the Sp was estimated to 0.98 (95% CPI: 0.95; 1.0), based on data from 368 goats.

Table 5. Sensitivity and specificity estimates from studies of interferon-γ tests for detection of infectious (I) and infected (E) cattle

Condition		Test	Anti-	Sample size[#]		Test outcome[¤]		Se[§]	Sp[§]	Age	Reference
Se[§]	Sp[§]		gen	C+	C-	T+|C+	T-|C-				
I	I	Various	PPDa	8	16	1	14	0.13	0.88	?	Paolicchi et al., 2003
I	E	Bovigam	PPDj	8	53	4	50	0.50	0.94	1-2 yrs	Huda et al., 2004
I	E	Bovigam	PPDj	13	65	11	50	0.85	0.94	2-3 yrs	Huda et al., 2004
I	E	Bovigam	PPDj	28	65	21	14	0.75	0.95	> 3 yrs	Huda et al., 2004

[§]) Se= Sensitivity; Sp=Specificity; I=Infectious; E=Infected.
[#]) Sample sizes for animals with condition (C+) and without condition (C-)
[¤]) No. of test-positive (T+) or test-negative (T-) given the condition (C)

6.6. Goats - Serum antibody ELISA

As in cattle, the most widely assessed test in goats has been the serum ELISA for detection of antibodies. However, contrary to cattle, the most widely condition detected has been affected and infected animals, whereas infectious animals has rarely been the condition detected (Table 6). Nonetheless, it appears that ELISA is more accurate in detection of all conditions than in cattle, with Se ranging from 0.82 to 1.0 for affected animals and 0.63–0.84 for infected animals. The corresponding Sp generally range from 0.92 to 1.0, although in one study with a very high Se of 0.91 the corresponding Sp was 0.79 only (Dimareli-Malli et al., 2004).

Table 6. Sensitivity and specificity estimates from studies of serum antibody ELISAs for detection of affected (A), infectious (I) and infected (E) goats

Condition		Test	Antigen	Sample size[#]		Test outcome[¤]		Se[§]	Sp[§]	Age	Reference
Se[§]	Sp[§]			C+	C-	T+|C+	T-|C-				
A	A	HerdChek	IDEXX	44	62	36	59	0.82	0.95	?	Dimareli-Malli et al., 2004
A	A	Various	Various	44	62	40	49	0.91	0.79	?	Dimareli-Malli et al., 2004
A	A	Various	Various	44	62	38	57	0.86	0.92	?	Dimareli-Malli et al., 2004
A	A	Various	Various	44	62	40	57	0.91	0.92	?	Dimareli-Malli et al., 2004
A	A	Parachek	VRI316	15	11	13	11	0.87	1.00	?	Milner et al., 1989
A	I	Various	Various	16	63	14	59	0.88	0.94	?	Molina et al., 1991
A	I	Various	PPA3	17	63	15	60	0.88	0.95	?	Molina et al., 1991
A	E	Various	PPA3	35	61	35	57	1.00	0.93	?	Molina Cabellero et al., 1993
I	E	HerdChek	IDEXX	35	123	19	123	0.54	1.00	?	Burnside & Rowley, 1994
I	NA	Various	Various	36		33		0.92		?	Tripathi et al., 2006
E	E	Pourquier	Various	36	945	28	945	0.78	1.00	> 1 yr	Gumber et al., 2006
E	E	HerdChek	IDEXX	368 combined				0.63	0.95	> 1 yr	Kostoulas et al., 2006
E	E	Parachek	VRI316	19	1000	16	997	0.84	1.00	?	Whittington et al., 2003
E	E	HerdChek	IDEXX	47	1000	39	995	0.83	1.00	?	Whittington et al., 2003

[§]) Se= Sensitivity; Sp=Specificity; A=Affected; I=Infectious; E=Infected; NA=Not available; [#]) Sample sizes for animals with condition (C+) and without condition (C-); [¤]) No. of test-positive (T+) or test-negative (T-) given the condition (C)

6.7. Sheep – Faecal Culture

FC has only been evaluated for the condition infected in sheep, and in one study only. The Se was estimated to 0.16 (95% CPI: 0.02; 0.48) and the Sp to 0.97 (95% CPI: 0.95–0.99) based on data from 368 animals evaluated in a latent class analysis (Kostoulas et al., 2006).

6.8. Sheep - Serum antibody ELISA

Ten serum antibody ELISAs have been evaluated for use in sheep (Table 7). In the evaluation of 5 tests, the Se were evaluated for detection of affected sheep and in the evaluation of 5 tests, infected animals were the target condition. All studies were based on detection of non-infected animals for estimation of Sp. The most widely used test was Parachek, and in 7 of the 10 studies, the antigen used was apparently MAP VRI316. For affected animals, the Se varied from 0.36 to 0.85, and for infected animals, the Se were 0.16–0.44. Sp ranged from 0.95 to 0.99 (Table 7).

Table 7. Sensitivity and specificity estimates from studies of serum antibody ELISAs for detection of affected (A), infectious (I) and infected (E) sheep

Condition		Test	Antigen	Sample size[#]		Test outcome[¤]		Se[§]	Sp[§]	Age	Reference
Se[§]	Sp[§]			C+	C-	T+|C+	T-|C-				
A	E	Parachek	VRI316	32	43	20	42	0.63	0.98	2-5 yr	Clarke et al., 1996
A	E	Various	Various	12	10	10	10	0.83	1.0	2-4 yr	Gwozdz et al., 1997
A	E	Parachek	VRI316	12	10	10	10	0.83	1.0	?	Gwozdz et al., 1997
A	E	Parachek	VRI316	59	253	50	252	0.85	1.0	?	Hilbink et al., 1994
A	E	Various	Various	59	252	21	247	0.36	0.98	?	Hilbink et al., 1994
E	E	Parachek	VRI316	120	1137	53	1125	0.44	0.99	> 1 yr	Hope et al., 2000
E	E	Various	Various	368 combined				0.37	0.97	> 1 yr	Kostoulas et al., 2006
E	E	Parachek	VRI316	2465 combined				0.16	0.98	> 1 yr	Robbe-Austerman et al., 2006
E	E	Various	VRI316	224	1748	93	1661	0.42	0.95	> 1 yr	Sergeant et al., 2003
E	E	Various	VRI316	224	1748	49	1731	0.22	0.99	> 1 yr	Sergeant et al., 2003

[§]) Se= Sensitivity; Sp=Specificity; A=Affected; I=Infectious; E=Infected; NA=Not available.

[#]) Sample sizes for animals with condition (C+) and without condition (C-)

[¤]) No. of test-positive (T+) or test-negative (T-) given the condition (C)

6.9. Deer – Serum antibody ELISA

In clinically affected deer, two non-commercial serum antibody ELISAs have been evaluated (Griffin et al., 2005). The ELISA based on PPDj antigen from CIDC (Lelystad, The Netherlands) had a Se of 0.83 (95% CI: 0.74; 0.89) (among 102 clinically affected deer) and a Sp of 1.00 (exact 95% CI: 0.9906; 0.9998) among animals from herds with no MAP infection history or clinical signs associated with MAP. An ELISA based on PPA3 antigen from Allied Monitor (Fayette, MO, USA) had a Se of 0.85 (95% CI: 0.77; 0.91) with a Sp of 0.98 (95% CI: 0.96; 0.99).

6.10. Overall summary of sensitivity and specificity

An overall summary of Se and Sp for each animal species, test and target condition is shown in Table 8, based on information in Sections Sections 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8 and 6.9. Both FC and ELISA generally have medium to high Se for detection of affected and infectious adult cattle, with Sp of 1.0 by definition. The range of Se of ELISA for detection of infectious cattle is huge, which is partly a reflection of the number of test-evaluations included in this group. The Se of FC for detection of infected cattle are in the range 0.23–0.29, which may be slightly better than Se of ELISA (range 0.07–0.39). However, a given Se

of ELISA needs to be evaluated in combination with Sp, and it cannot be determined which of the tests that are the better. The variation in test-accuracy within test-evaluations for IFN-γ is huge, but basically only two studies have been performed, with significant differences in the results. It cannot be determined which of the results that are the more reliable.

Table 8. Summary (range) of reported sensitivities and specificities of faecal culture (FC), serum antibody ELISA (SELISA), milk antibody ELISA (MELISA) and interferon-γ tests for diagnosis of three stages of infection with *Mycobacterium avium* subsp. *paratuberculosis* in cattle, deer, goats and sheep.

	Sensitivities (range)				Specificities (range)			
	FC	SELISA	MELISA	IFN-γ	FC	SELISA	MELISA	IFN-γ
Cattle								
Affected	0.70[#1]	0.50; 0.87[#3]	None	None	1.0[§]	None	None	None
Infectious	0.74[#1]	0.24; 0.94[#30]	0.21; 0.61[#5]	0.13; 0.85[#4]	1.0[§]	0.40; 1.0[#15]	0.83; 0.99[#2]	0.88[#1]
Infected	0.23;0.29[3#]	0.07; 0.22[#5]	0.39[#1]	None	0.98[#1]	0.85; 1.0[#19]	0.96; 1.0[#2]	0.94; 0.95[#3]
Deer								
Affected	None	0.83; 0.85[1#]	None	None	1.0[§]	0.98; 1.0[2#]	None	None
Infectious	None	None	None	None	1.0[§]	None	None	None
Infected	None	None	None	None	None	None	None	None
Goats								
Affected	None	0.82; 1.0[#8]	None	None	1.0[§]	0.79; 1.0[#5]	None	None
Infectious	None	0.54; 0.92[#2]	None	None	1.0[§]	0.94; 0.95[#2]	None	None
Infected	0.08[1#]	0.63; 0.84[#4]	None	None	0.98[#1]	0.93; 1.0[#6]	None	None
Sheep								
Affected	None	0.36; 0.85[#5]	None	None	1.0[§]	None	None	None
Infectious	None	None	None	None	1.0[§]	None	None	None
Infected	0.16[#1]	0.16; 0.44[#5]	None	None	0.97[#1]	0.95; 1.0[#10]	None	None

[#]) Number given is the number of test-evaluations included in the summary. One study (Billman-Jacobe et al., 1992) was excluded from group FC, Infected, Cattle because of selection procedure of animals.
[§]) Specificity is 1.0 by definition

In deer, only one study including two ELISAs used for clinically affected animals have been reported, with promising results. However, the lack of studies on infected and infectious deer prompts for studies on these conditions.

ELISA used for infected, infectious and affected goats indicates that this test can have utility because of generally high Se, irrespective of target condition. The Se of FC for detection of infected goats is, however, not promising with a Se of 0.08 and a Sp of 0.98, in the one study reported. Further studies are needed to draw reasonable conclusions. However, the Se of FC is comparable to that of sheep, indicating that the estimate may be valid for the particular test. The FC method evaluated for both sheep and goats originate from the same study, and other FC methods may prove to be superior. IFN-γ and milk ELISA have not been evaluated in deer, goats and sheep. Results of test-evaluations for serum ELISA for sheep show considerable variation for both affected and infected animals, but the Se for serum ELISAs used for infected animals appear to be lower than serum ELISAs used for infected animals, as would be expected.

7. DISCUSSION

This report summarises Se and Sp estimates for detection of animals infected by, infectious with or affected by MAP, obtained through a critical review of literature. The report is the first to divide animals into the three target conditions (infected, infectious and affected), but these can be very useful in the processes where decision makers have to make decisions related

to different conditions. The division also appears to provide estimates which are more homogenous for a given test than had the estimates been reported for one group only. As an example, the Sp of FC were 0.96, 0.98 and 0.97, for cattle, goats and sheep respectively. Serum antibody ELISA used to detect infected cattle had Se in the range from 0.07 to 0.22, whereas serum antibody ELISAs for detection of infected goats were in the range 0.63–0.84. These ranges are much narrower than could be expected. The narrow range is of course partly a function of the low number of test-evaluations per group, but it still appears to be narrower irrespective hereof.

Division of the target conditions into affected, infectious and infected animals to some extent reduced the variation of Se within a group, as would be expected. For a decision maker, estimates of Se and Sp that are specific for a given target condition, should be preferred, because the decision maker can then report the probability of having the given condition based on the test result. If no distinction between target conditions is made, a Se and Sp is for an average of infected, infectious and affected animals will have to be assumed. In Collins et al. (2006), such an average must have been assumed, as we were unable to identify in literature the Se and Sp estimates reported as “assumptions for test Se and Sp”. As an example, they reported a Se of 0.60 ± 0.05 for FC in cattle, given the best test is used. This figure is twice the size of the Se reported in literature for detection of infected animals. The Sp of FC was reported to 0.999 ± 0.001, which would be applicable only for non-infected animals in non-infected herds, not a randomly selected animal. The estimates from literature suggest the Sp to be 0.98 for cattle, which is supported by the estimates of 0.97 and 0.98 for sheep and goats, respectively. The target conditions and basis for the accuracies reported by Collins et al. (2006) as “consensus recommendations” were not given. Differences in target condition and the choice of “best test” could be the reason for differences between their figures and the estimates reported in literature.

Combining the estimates within groups of animals and conditions into one estimate with associated uncertainty estimates would have been preferred. With the current approach, it is also problematic that there is a uniform weighting of test-evaluations irrespective of sample size and year of publication. The latter is due to potential improvement of tests with time. A formal meta-analysis was not performed, primarily due to the differences in test protocols used, particularly differences in antigen formulations and chosen cut-offs, making it hard to justify comparisons across studies.

The numbers given in the present report may be useful not only to decision makers, but also as input parameters in simulation studies. There are still a number of test-target condition combinations, which have not been evaluated, particularly in deer. There is therefore still a need for further well-planned diagnostic test-evaluations. Also, the quality of test-evaluations was inadequate for across-study comparisons also because very few studies report the target and the study population. Improvements in planning, conducting and reporting test-evaluations are generally required, not only for test evaluations related to MAP infection. It is recommended to follow the guidelines given by Greiner and Gardner (2000). The ideal test-evaluation for a set of diagnostic paratuberculosis data have still to be published for several reasons. The complicated nature of the long, slowly developing MAP infections and the lack of good reference tests will most likely introduce selection bias in traditional test-evaluations. Therefore, while it is relatively simple to include covariate effects such as animal or laboratory effects using traditional methods, the resulting accuracy estimates will most likely be biased. As an alternative, multivariable test-evaluations without

selection bias can be performed use of some latent class methods such as Bayesian analyses. However, such methods are not always as computationally simple to perform, and the interpretation of target conditions is not always straight forward when these models are used. For the target condition “infection”, latent-class models are probably the best alternative, whereas for conditions “infectious” and “affected”, traditional methods may be better used, because these conditions generally are easier detected. An ‘ideal’ test-evaluation could include a longitudinal study design over the entire lifetime of the animals studied. The animals would have to be tested regularly with an agent detecting test (e.g. FC), a test detecting cell-mediated immune responses (e.g. IFN-γ) and antibody reactions (e.g. serum ELISA), ultimately ending with a post-mortem histopathological evaluation of up to 100 tissues per animal to determine the infection status of the animal. Such a study would be extremely expensive and perhaps not even provide the necessary information. However, longitudinal studies and/or more frequent use of latent class methodology could provide better test-evaluations than is currently seen. Latent class models with inclusion of covariates have been used to demonstrate the improved accuracy obtained when several tests are used, i.e. reduced milk production combined with FC and milk antibody ELISA (Wang et al., 2006). Such an approach would reduce the need for division into different conditions as suggested here. The approach needs to be developed further to include repeated test data over time. At the current stage, division into detection of different conditions will however be beneficial, particularly for decision makers.

8. CONCLUSION

The Se and Sp of diagnostic tests for various stages of MAP infections varied significantly, but formal comparison of the different tests cannot be justified. The main reasons are variations in study design, test components and target conditions. Stratification of target conditions into those relevant for decision makers can decrease the variation of each test in each animal species, thereby improving the interpretation. However, the accuracies of the various test types appear to vary from species to species for different target definitions, therefore interpretation of diagnostic test information should be made by species, target condition and test. There is still a profound lack of reliable test evaluations, and future assessments should be conducted more stringently to allow appropriate interpretation and comparison across populations.

ACKNOWLEDGEMENTS

This study was co-funded by the European Commission within the Sixth Framework Programme, as part of the project ParaTBTools (contract no. 023106 (FOOD)).

REFERENCES

Abbas, B., Riemann, H.P., Lonnerdal, B., 1983. Isolation of specific peptides from *Mycobacterium paratuberculosis* protoplasm and their use in an enzyme-linked immunosorbent assay for detection of paratuberculosis (Johne’s disease) in cattle. Am. J. Vet. Res. 44, 2229-2236.

Bech-Nielsen, S., Jorgensen, J.B., Ahrens, P., Feld, N.C., 1992. Diagnostic accuracy of a *Mycobacterium phlei*-absorbed serum enzyme-linked immunosorbent assay for diagnosis of bovine paratuberculosis in dairy cows. J. Clin. Microbiol. 30, 613-618.

Benedictus, G., Dijkhuizen, A.A., Stelwagen, J., 1987. Economic losses due to paratuberculosis in dairy cattle. Vet. Rec. 121, 142-146.

Berghaus, R.D., Farver, T.B., Anderson, R.J., Adaska, J.M., Gardner, I.A., 2006. Use of age and milk production data to improve the ability of enzyme-linked immunosorbent assay test results to predict *Mycobacterium avium* ssp. *paratuberculosis* fecal culture status. J. Vet. Diagn. Invest. 18, 233-242.

Billman-Jacobe, H., Carrigan, M., Cockram, F., Corner, L.A., Gill, I.J., Hill, J.F., Jessep, T., Milner, A.R., Wood, P.R., 1992. A comparison of the interferon gamma assay with the absorbed ELISA for the diagnosis of Johne's disease in cattle. Aust. Vet. J. 69, 25-28.

Burnside, D.M., Rowley, B.O., 1994. Evaluation of an enzyme-linked immunosorbent assay for diagnosis of paratuberculosis in goats. Am. J. Vet. Res. 55, 465-466.

Cho D., Collins, M.T., 2006. Comparison of proteosomes and antigenicities of secreted and cellular proteins produced by *Mycobacterium paratuberculosis*. Clin. Vacc. Immunol. 13, 1155-1161.

Chiodini, R.J., 1993. Abolish *Mycobacterium paratuberculosis* Strain 18. J .Clin. Microbiol. 31, 1956-1958.

Clarke, C.J., Patterson, I.A., Armstrong, K.E., Low, J.C., 1996. Comparison of the absorbed ELISA and agar gel immunodiffusion test with clinicopathological findings in ovine clinical paratuberculosis. Vet. Rec. 139, 618-621.

Colgrove, G.S., Thoen, C.O., Blackburn, B.O., Murphy, C.D., 1989. Paratuberculosis in cattle: a comparison of three serologic tests with results of fecal culture. Vet. Microbiol. 19, 183-187.

Collins, M.T., Sockett, D.C., Ridge, S., Cox, J.C., 1991. Evaluation of a commercial enzyme-linked immunosorbent assay for Johne's disease. J. Clin. Microbiol. 29, 272-276.

Collins, M.T., Wells, S.J., Petrini, K.R., Collins, J.E., Schultz, R.D., Whitlock, R.H., 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab Immunol. 12, 685-692.

Collins, M.T., Gardner, I.A., Garry, F.B., Roussel, A.J., Wells, S.J., 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229, 1912-1919.

Dimareli-Malli, Z., Samarineanu, M., Sarca, M., Zintzaras, E., Sarris, K., Tsitsamis, S., 2004. Statistical evaluation of ELISA methods for testing caprine paratuberculosis. Int. J. Appl. Res. Vet. Med. 2, 10-16.

Doyle, T.M., 1953. Susceptibility to Johne's disease in relation to age. Vet. Rec. 65, 363-365.

Doyle, T.M., Spears, H.N., 1951. A Johne's disease survey. Vet. Rec. 63, 355-359.

Eamens, G.J., Whittington, R.J., Marsh, I.B., Turner, M.J., Saunders, V., Kemsley, P.D., Rayward, D., 2000. Comparative sensitivity of various faecal culture methods and ELISA in dairy cattle herds with endemic Johne's disease. Vet. Microbiol. 77, 357-367.

Eda, S., Bannantine, J.P., Waters, W.R., Mori Y., Whitlock, R.H., Scott, M.C., Speer, C.A., 2006. A highly sensitive and subspecies-specific surface antigen enzyme-linked immunosorbent assay for diagnosis of Johne's disease. Clin. Vaccine Immunol. 13, 837-844.

Egan, J., Weavers, E., O'Grady, D., 1999. An evaluation of diagnostic tests for Johne's disease in cattle. Irish Vet. J. 52, 86-89.

Glanemann, B., Hoelzle, L.E., Bogli-Stuber, K., Jemmi, T., Wittenbrink, M.M., 2004. Detection of *Mycobacterium avium* subspecies *paratuberculosis* in Swiss dairy cattle by culture and serology. Schweiz. Arch. Tierheilkd. 146, 409-415.

Greiner, M., Gardner, I.A., 2000. Epidemiologic issues in the validation of veterinary diagnostic tests. Prev. Vet. Med. 45, 3-22.

Griffin, J.F., Spittle, E., Rodgers, C.R., Liggett, S., Cooper, M., Bakker, D., Bannantine, J.P., 2005. Immunoglobulin G1 enzyme-linked immunosorbent assay for diagnosis of Johne's Disease in red deer (*Cervus elaphus*). Clin. Diagn. Lab Immunol. 12, 1401-1409.

Gumber, S., Eamens, G., Whittington, R.J., 2006. Evaluation of a Pourquier ELISA kit in relation to agar gel immunodiffusion (AGID) test for assessment of the humoral immune response in sheep and goats with and without *Mycobacterium paratuberculosis* infection. Vet. Microbiol. 115, 91-101.

Gwozdz, J.M., Reichel, M.P., Murray, A., Manktelow, W., West, D.M., Thompson, K.G., 1997. Detection of *Mycobacterium avium* subsp. *paratuberculosis* in ovine tissues and blood by the polymerase chain reaction. Vet. Microbiol. 57, 233-244.

Hendrick, S.H., Duffield, T.E., Kelton, D.E., Leslie, K.E., Lissemore, K.D., Archambault, M., 2005. Evaluation of enzyme-linked immunosorbent assays performed on milk and serum samples for detection of paratuberculosis in lactating dairy cows. J. Am. Vet. Med. Assoc. 226, 424-428.

Hilbink, F., West, D.M., de Lisle, G.W., Kittelberger, R., Hosie, B.D., Hutton, J., Cooke, M.M., Penrose, M., 1994. Comparison of a complement fixation test, a gel diffusion test and two absorbed and unabsorbed ELISAs for the diagnosis of paratuberculosis in sheep. Vet. Microbiol. 41, 107-116.

Holmes, I.R., Jubb, T.F., Callinan, A.P., 2004. Infection rates in reactors to an absorbed ELISA used in a test and cull program for bovine Johne's disease. Aust. Vet. J. 82, 233-235.

Hope, A.F., Kluver, P.F., Jones, S.L., Condron, R.J., 2000. Sensitivity and specificity of two serological tests for the detection of ovine paratuberculosis. Aust. Vet. J. 78, 850-856.

Houe, H., Ersbøll, A.K., Toft, N. (Eds): Introduction to Veterinary Epidemiology. Biofolia, Frederiksberg, Denmark, pp. 133-151.

Huda, A., Jungersen, G., Lind, P., 2004. Longitudinal study of interferon-gamma, serum antibody and milk antibody responses in cattle infected with *Mycobacterium avium* subsp. *paratuberculosis*. Vet. Microbiol. 104, 43-53.

Jubb, T.F., Galvin, J.W., 2004. Effect of a test and control program for Johne's disease in Victorian beef herds 1992-2002. Aust. Vet. J. 82, 164-166.

Jubb, T.F., Sergeant, E.S., Callinan, A.P., Galvin, J., 2004. Estimate of the sensitivity of an ELISA used to detect Johne's disease in Victorian dairy cattle herds. Aust. Vet. J. 82, 569-573.

Kalis, C.H., Barkema, H.W., Hesselink, J.W., van Maanen, C., Collins, M.T., 2002. Evaluation of two absorbed enzyme-linked immunosorbent assays and a complement fixation test as replacements for fecal culture in the detection of cows shedding *Mycobacterium avium* subspecies *paratuberculosis*. J. Vet. Diagn. Invest. 14, 219-224.

Kennedy, D.J., Benedictus, G., 2001. Control of *Mycobacterium avium* subsp. *paratuberculosis* infection in agricultural species. Rev. sci. tech. Off. Int. Epiz. 20, 151-179.

Klausen, J., Huda, A., Ekeroth, L., Ahrens, P., 2003. Evaluation of serum and milk ELISAs for paratuberculosis in Danish dairy cattle. Prev. Vet. Med. 58, 171-178.

Koets, A.P. , Rutten, V.P.M.G., de Boer, M., Bakker, D., Valentin-Weigand, P., van Eden, W., 2001. Differential changes in heat shock protein-, lipoarabinomannan-, and purified protein derivative-specific immunoglobulin G1 and G2 isotype responses during bovine *Mycobacterium avium* subsp. *paratuberculosis* infection. Infect. Immun. 69, 1492-1498.

Kostoulas, P., Leontides, L., Enoe, C., Billinis, C., Florou, M., Sofia, M., 2006. Bayesian estimation of sensitivity and specificity of serum ELISA and faecal culture for diagnosis of paratuberculosis in Greek dairy sheep and goats. Prev. Vet. Med. 76, 56-73.

Kudahl, A.B., Nielsen, S.S., Sørensen, J.T., 2004. Relationship between antibodies against *Mycobacterium avium* subsp. *paratuberculosis* in milk and shape of lactation curves. Prev. Vet. Med. 62, 119-134.

Kudahl, A.B., Østergaard, S., Sørensen, J.T., Nielsen, S.S., 2007. A stochastic model simulating paratuberculosis in a dairy herd. Prev. Vet. Med. 78, 97-117.

Lepper, A.W.D., Wilks, C.R., Kotiw, M., Whitehead, J.T., Swart, K.S., 1989. Sequential bacteriologic observations in relation to cell-mediated and humoral antibody responses of cattle infected with *Mycobacterium paratuberculosis* and maintained on normal or high iron intake. Aust. Vet. J. 66, 50-55

Lombard, J.E., Byrem, T.M., Wagner, B.A., McCluskey, B.J., 2006. Comparison of milk and serum enzyme-linked immunosorbent assays for diagnosis of *Mycobacterium avium* subspecies *paratuberculosis* infection in dairy cattle. J. Vet. Diagn. Invest. 18, 448-458.

Martin, S.W., Meek, A.H., Willeberg, P., 1987. Veterinary Epidemiology, chap. 10, p. 245-258. Iowa State University Press, Ames, Iowa, USA.

McKenna, S.L., Keefe, G.P., Barkema, H.W., Sockett, D.C., 2005. Evaluation of three ELISAs for *Mycobacterium avium* subsp. *paratuberculosis* using tissue and fecal culture as comparison standards. Vet. Microbiol. 110, 105-111.

McNab, W.B., Meek, A.H., Duncan, J.R., Brooks, B.W., Van Dreumel, A.A., Martin, S.W., Nielsen, K.H., Sugden, E.A., Turcotte, C., 1991. An evaluation of selected screening tests for bovine paratuberculosis. Can. J. Vet. Res. 55, 252-259.

Milner, A.R., Mack, W.N., Coates, K.J.. 1989. A modified ELISA for the detection of goats infected with *Mycobacterium paratuberculosis*. Aust. Vet. J. 66, 305-307.

Milner, A.R., Mack, W.N., Coates, K.J., Hill, J., Gill, I., Sheldrick, P., 1990. The sensitivity and specificity of a modified ELISA for the diagnosis of Johne's disease from a field trial in cattle. Vet. Microbiol. 25, 193-198.

Molina, A., Morera, L., Llanes, D., 1991. Enzyme-linked immunosorbent assay for detection of antibodies against *Mycobacterium paratuberculosis* in goats. Am. J. Vet. Res. 52, 863-868.

Molina Caballero, J.M., Anguiano, A., Ferrer, O., Serrano, E., Uceda, A., 1993. Use of an enzyme-linked immunosorbent assay for serodiagnosis of clinical paratuberculosis in goats. Study by Western blotting of false-positive reactions. Rev. Sci. Tech. Off. Int. Epiz. 12, 629-638.

Mortensen, H., Nielsen, S.S., Berg, P., 2004. Genetic variation and heritability of the antibody response to *Mycobacterium avium* subsp. *paratuberculosis* in Danish Holstein cows. J. Dairy Sci. 87, 2108-2113.

Nielsen, S.S., Ersbøll, A.K., 2006. Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows. J. Dairy Sci. 89, 4557-4566.

Nielsen, S.S., Toft, N., 2006a. Age-specific characteristics of ELISA and fecal culture for purpose specific testing for paratuberculosis. J. Dairy Sci. 89, 569-579.

Nielsen, S.S., Toft, N., 2006b. Purpose-related interpretation of diagnostic test-information used for bovine paratuberculosis. Proceedings of the 11th International Symposium on Veterinary Epidemiology and Economics, 6-11 August 2006. Cairns, Australia, p. 620-622.

Nielsen, S.S., Houe, H., Thamsborg, S.M., Bitsch, V., 2001. Comparison of two enzyme-linked immunosorbent assays for serologic diagnosis of paratuberculosis (Johne's disease) in cattle using different subspecies strains of *Mycobacterium avium*. J. Vet. Diagn. Invest. 13, 164-166.

Nielsen, S.S., Grønbæk, C., Agger, J.F., Houe, H., 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53, 191-204.

Osterstock, J.B., Fosgate, G.T., Norby, B., Manning, E.J.B., Collins, M.T., Roussel, A.J., 2007. Contribution of environmental mycobacteria to false-positive serum ELISA results for paratuberculosis. J. Am. Vet. Med. Assoc., 230, 896-901.

Paolicchi, F.A., Zumarraga, M.J., Gioffre, A., Zamorano, P., Morsella, C., Verna, A., Cataldi, A., Alito, A., Romano, M., 2003. Application of different methods for the diagnosis of paratuberculosis in a dairy cattle herd in Argentina. J. Vet. Med. B 50, 20-26.

Pitt, D.J., Pinch, D.S., Janmaat, A., Condron, R.J., 2002. An estimate of specificity for a Johne's disease absorbed ELISA in northern Australian cattle. Aust. Vet. J. 80, 57-60.

Reichel, M.P., Kittelberger, R., Penrose, M.E., Meynell, R.M., Cousins, D., Ellis, T., Mutharia, L.M., Sugden, E.A., Johns, A.H., de Lisle, G.W., 1999. Comparison of serological tests and faecal culture for the detection of *Mycobacterium avium* subsp. *paratuberculosis* infection in cattle and analysis of the antigens involved. Vet. Microbiol. 66, 135-150.

Robbe-Austerman, S., Gardner, I.A., Thomsen, B.V., Morrical, D.G., Martin, B.M., Palmer, M.V., Thoen, C.O., Ewing, C., 2006. Sensitivity and specificity of the agar-gel-immunodiffusion test, ELISA and the skin test for detection of paratuberculosis in United States Midwest sheep populations. Vet. Res. 37, 553-564.

Roussel, A.J., Fosgate, G.T., Manning, E.J.B., Collins, M.T., 2007. Association of fecal shedding of mycobacteria with high ELISA-determined seroprevalence for paratuberculosis in beef herds. J. Am. Vet. Med. Assoc., 230, 890-895.

Sergeant, E.S., Marshall, D.J., Eamens, G.J., Kearns, C., Whittington, R.J., 2003. Evaluation of an absorbed ELISA and an agar-gel immuno-diffusion test for ovine paratuberculosis in sheep in Australia. Prev. Vet. Med. 61, 235-248.

Sockett, D.C., Conrad, T.A., Thomas, C.B., Collins, M.T., 1992a. Evaluation of four serological tests for bovine paratuberculosis. J. Clin. Microbiol. 30,1134-1139.

Sockett, D.C., Carr, D.J., Collins, M.T., 1992b. Evaluation of conventional and radiometric fecal culture and a commercial DNA probe for diagnosis of *Mycobacterium paratuberculosis* infections in cattle. Can. J. Vet. Res. 56, 148-153.

Spangler, E., Bech-Nielsen, S., Heider, L.E., 1992. Diagnostic performance of two serologic tests and fecal culture for subclinical paratuberculosis, and associations with production. Prev. Vet. Med. 13, 185-195.

Stabel, J.R., 2000. Transistions in immune responses to *Mycobacterium paratuberculosis*. Vet. Microbiol. 77, 465-473.

Stabel, J.R., Wells, S.J., Wagner, B.A., 2002. Relationships between fecal culture, ELISA, and bulk tank milk test results for Johne's disease in US dairy herds. J. Dairy Sci. 85, 525-531.

Sweeney, R.W., Whitlock, R.H., Hamir, A.N., Rosenberger, A.E., Herr, S.A., 1992a. Isolation of *Mycobacterium paratuberculosis* after oral inoculation in uninfected animals. Am. J. Vet. Res. 53, 1312-1314.

Sweeney, R.W., Whitlock, R.H., Buckley, C.L., Spencer, P., Rosenberger, A.E., Hutchinson, L.J., 1994. Diagnosis of paratuberculosis in dairy cattle, using enzyme-linked immunosorbent assay for detection of antibodies against *Mycobacterium paratuberculosis* in milk. Am. J. Vet. Res. 55, 905-909.

Sweeney, R.W., Whitlock, R.H., Buckley, C.L., Spencer, P.A., 1995. Evaluation of a commercial enzyme-linked immunosorbent assay for the diagnosis of paratuberculosis in dairy cattle. J. Vet. Diagn. Invest. 7, 488-493.

Taylor, A.W., 1953. Experimental Johne's disease in cattle. J. Comp. Pathol. 63:355–367.

Tripathi, B.N., Periasamy, S., Paliwal, O.P., Singh, N., 2006. Comparison of IS*900* tissue PCR, bacterial culture, johnin and serological tests for diagnosis of naturally occurring paratuberculosis in goats. Vet. Microbiol. 116, 129-137.

Wang, C., Turnbull, B.W., Grohn, Y.T., Nielsen, S.S., 2006. Estimating receiver operating characteristic curves with covariates when there is no perfect reference test for diagnosis of Johne's disease. J. Dairy Sci. 89, 3038-3046.

Waters, W.R., Miller J.M., Palmer M.V., Stabel J.R., Jones D.E., Koistinen K.A., Steadham E.M., Hamilton M.J., Davis W.C., Bannantine J.P., 2003. Early induction of humoral and cellular immune responses during experimental *Mycobacterium avium* subsp. *paratuberculosis* infection of calves. Infect Immun. 71, 5130-5138.

Whitlock, R.H., Rosenberger, A.E., Sweeney, R.W., Spencer, P.A., 1996. Distribution of *M. paratuberculosis* in tissues of cattle from herds infected with Johne's disease. *In*: Chiodini, R.J., Hines II, M.E., Collins, M.T. (eds): Proceedings of the Fifth International Colloquium on Paratuberculosis, Madison, Wisconsin, USA, Sept. 29 – Oct. 4, 1996, p. 168-174.

Whitlock, R.H., Wells, S.J., Sweeney, R.W., Van Tiem, J., 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): sensitivity and specificity of each method. Vet. Microbiol. 77, 387-398.

Whittington, R.J., Eamens, G.J., Cousins, D.V., 2003. Specificity of absorbed ELISA and agar gel immuno-diffusion tests for paratuberculosis in goats with observations about use of these tests in infected goats. Aust. Vet. J. 81, 71-75.

Yekutiel, P., 1980. Eradication of infectious diseases: A critical study. *In*: Contribution to Epidemiology and Biostatistics, vol. 2, Karger, Basel, Switzerland, 164 pp.

Yokomizo, Y., Kishima, M., Mori, Y., Nishimori, K., 1991. Evaluation of enzyme-linked immunosorbent assay in comparison with complement fixation test for the diagnosis of subclinical paratuberculosis in cattle. J. Vet. Med. Sci. 53, 577-584.

Paper II

Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows

Nielsen SS, Ersbøll AK

Published in
Journal of Dairy Science 2006, 89, 4557–4566

Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows

S. S. Nielsen, and A. K. Ersbøll

Department of Large Animal Sciences, The Royal Veterinary and Agricultural University, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark

ABSTRACT

Paratuberculosis is a chronic infection of ruminants and other species caused by *Mycobacterium avium* ssp. *paratuberculosis* (Map). Establishing test strategies for paratuberculosis will require insight into the temporal aspects of certainty with a given test. In this study, the age at which cows tested positive by ELISA and fecal culture (FC) was investigated by use of time-to-event analyses. The effects of herd, parity, and shedding group were evaluated at the age of test-positive ELISA and FC, respectively. Finally, the test frequency was investigated for the probability of cows being tested ELISA-positive. Milk and fecal samples were collected repeatedly over a 3-yr period from 1,776 Danish dairy cows from 8 herds. The milk samples were tested for the presence of antibodies by using an ELISA, and an FC test was used for detection of Map. Repeated ELISA testing detected 98 and 95% of cows classified as high and low shedders, respectively, suggesting that most infected cows develop antibodies. Among the high shedders, 50% were positive before 4.3 yr of age (quartiles 1 to 3: 3.4 to 5.7 yr of age). Repeated FC detected only 72% of the cows that were ELISA-positive, and 50% of the ELISA-positive cows were detected by FC at 7.6 yr of age. The age with the highest probability of testing positive was determined as the interval with the steepest slope in the survival probability plots. The highest probability of testing positive by ELISA was from 2.5 to 4.5 yr of age. The highest probability of testing positive by FC was from 2.5 to 5.5 yr of age. For both ELISA and FC, testing positive was highest in the first 300 d in milk. For cows younger than 4 yr of age, monthly testing with ELISA, compared with testing every 2 yr, could increase the probability of detecting cows with antibodies by 19%. In older cows, there were no apparent differences in the probability of testing positive by monthly sampling compared with sampling every second year. Therefore, for older animals the effect of more frequent sampling would be for early detection rather than to obtain additional information. Cows shedding high numbers of Map will produce antibodies, although not necessarily concomitantly with the shedding. These antibodies can be detected by ELISA with a test strategy that is different for younger and older cows. We suggest testing younger cows more frequently than older cows and that testing should be done prior to 350 d in milk.

INTRODUCTION

Paratuberculosis in cattle is a chronic infection caused by *Mycobacterium avium* ssp. *paratuberculosis* (Map; Chiodini et al., 1984). Diagnosis is usually made by detection of Map by fecal culture (FC), or by detection of antibodies (indicating immune responses) by ELISA. Detection by FC requires that Map be shed in numbers high enough to elicit a positive culture, and ELISA requires that animals have antibodies (i.e., Map is seroconverted). The infection is initially controlled by a predominating T helper 1 (Th1) response, whereas loss of control occurs via a predominant Th2 response (Stabel, 2000). During the Th1 response,

Map is shed in small numbers, which may be sufficient to elicit a positive FC. The pattern of bacterial shedding is not well characterized in naturally infected animals for the various phases of the infection. Experimental infections suggest that bacterial shedding decreases 10 to 14 mo after inoculation, but increases again later. Seroconversion is seen around 10 mo postinoculation (Lepper et al., 1989); however, experimental studies cannot be used as indicators of when bacterial shedding and seroconversion occur because of the variation in infective doses occurring during natural infections with Map. Studies using fixed dosages and known times of infection have resulted in great variation in the time to occurrence of FC-positivity and ELISA-positivity (Lepper et al., 1989). The temporal variation in pathogenetic events may be further enhanced if the size and number of dosages varies. Chiodini et al. (1984) reported that the immunological and infectious properties of the infection are not fully characterized, but stated that most cases occur between 3 and 5 yr of age.

The ability of a test to diagnose paratuberculosis will be affected by the age of an animal because of the chronic nature of the infection. The variation in age at seroconversion or bacterial shedding is a factor that can greatly affect the sensitivity of a test when determined in a cross-sectional study, and age-specific sensitivities of a test would have to be estimated. Therefore, it is important to know the age when seroconversion and bacterial shedding occur in infected animals if the test is expected to detect these 2 events. Test sensitivities in the range of 7% (McKenna et al., 2005) to 45% (Collins et al., 2005) were demonstrated in recent evaluations of a range of ELISA tests. Likely explanations for the low and variable sensitivities could be that antibodies are not present because of the distribution of stages of infection in the cross-sectional samples, or that the ELISA was unable to detect occurring antibodies. A report of sensitivities estimated, relative to stage of infection, a 15% sensitivity of ELISA in detecting low Map-shedding cows and an 87% sensitivity of the same ELISA in detecting clinical cases (Sweeney et al., 1995).

Jubb and Galvin (2000) described the mean age of detection of ELISA-positive cows as 5.6 yr, but with a range of 2 to 18 yr. Similarly, van Schaik et al. (2003) and Kalis et al. (1999) demonstrated that older animals are more likely to be positive by ELISA and FC, respectively. None of the studies included time-to-event analyses (also called survival analyses), which would be more appropriate to manage censored animals and skewed distributions of the time to an event. In survival analyses, the time until an event occurs is regarded as the outcome variable (Hosmer and Lemeshow, 1999). The event could be ELISA-positive and the time would be time to ELISA-positivity for an individual animal. Censored animals would not have an observed time to an event, because the event does not occur for these animals for different reasons (e.g., withdrawal or an insufficient period of time for the event to occur). Because time-to-event data are continuous, it would be obvious to perform an ANOVA or linear regression to evaluate the effect of risk factors on the time to an event. Moreover, the assumption of normally distributed residuals is often violated because the distribution of time to an event is often skewed.

Our primary objectives were to determine 1) the time from birth to test-positivity and 2) the DIM at test-positivity. Test-positivity was defined as a) the occurrence of antibodies, estimated by a milk ELISA response, and b) the occurrence of detectable bacterial shedding, estimated by a fecal sample response. Our secondary objectives were to determine the influence of herd, parity, and shedding group on the time from birth to being test-positive (using the milk ELISA response and bacterial shedding, respectively). Finally, we wished to determine the effect of sampling frequency on test-positivity using the milk ELISA response.

MATERIALS AND METHODS

Herds and Animals

Samples were obtained from 8 Danish dairy herds from February 2000 to March 2003. Herd sizes ranged from 66 to 260 cows from October 1, 1999, to September 30, 2000. Herd milk production was between 5,922 and 10,060 kg of FCM per cow. Additional information regarding the herds is available in Nielsen and Toft (2006). Milk samples were collected from lactating cows 11 times/yr through the Danish milk recording scheme, and fecal samples were collected 4 times/yr from all cows present in the herd at the date of sampling.

Diagnostic information was collected from 1,885 cows. Information on dates of birth, dates of calving, and breed was obtained from the Danish Cattle Database. Cows (n = 100) without ELISA and FC test results were excluded, as were 9 Red Danish cows. The breed distribution of the remaining 1,776 cows was 1,260 Danish Holsteins, 407 Danish Jerseys, and 109 cross-breds. Age was calculated as the difference between sampling date and date of birth. For some analyses, the cows were divided into parity groups. These groups were parities 1, 2, and 3, corresponding to the grouping in Nielsen et al. (2002). The distribution of milk samples from these cows was as follows: minimum: 1 sample/cow; 25th percentile: 6 samples/cow; median: 11 samples/cow; 75th percentile: 18 samples/cow; maximum: 30 samples/cow. The distribution of fecal samples per cow was as follows: minimum: 1 sample/cow; 25th percentile: 3 samples/cow; median: 4 samples/cow; 75th percentile: 7 samples/cow; maximum: 13 samples/cow.

Diagnostic procedures

Milk samples were analyzed for the presence of antibodies using an absorbed indirect milk ELISA, based on an M. avium antigen, as previously described (Nielsen, 2002). The resulting measure was a corrected optical density, which was obtained by subtracting the mean optical density of a set of negative controls from the sample optical density. An animal was defined as antibody-positive on the date the moving average of 2 consecutive measurements was greater than a corrected optical density of 0.3 and the minimum of these 2 measurements was greater than 0.1. This test interpretation should provide a minimum specificity ranging from 0.997 at 2 yr of age to 0.93 at 5 yr of age, assuming that 8 negative FC tests obtained over a 2-yr period could be used as a reference for noninfected animals (Nielsen and Toft, 2006).

Fecal samples were cultured on Löwenstein–Jensen medium through July 2002 and on Herrold's egg yolk medium beginning in August 2002 (Nielsen et al., 2004). The latter procedure appeared more sensitive (Nielsen et al., 2004), although this was not observed in a prospective study (Nielsen and Toft, 2006). All positive isolates were confirmed for the presence of the IS900 insertion sequence by PCR. For the FC analysis, the date on which Map was detected was the date the cow was considered positive. To evaluate the antibody response, FC-positive cows were divided into 2 shedding groups: FCHigh and FCLow, where FCHigh was culture-positive cows with high bacterial counts (>10 cfu/g) or many repeated FC-positive tests without any nonzero counts of bacteria in between. The remaining FC-positive cows were classified as FCLow. These shedding groups were the same as those used in Toft et al. (2005). A thorough descriptive analysis of test responses is given in Nielsen and Toft (2006).

Statistical Analysis

The primary outcome variable was "age at test-positivity," defined as days from birth to either date of being tested positive (event) or the end of the study period (censored observation) for test-negative animals. Test-positivity was defined as antibody-positivity (ELISA) or shedding level (FC). The second outcome variable was "DIM at test-positivity," which was used to determine whether the test-positive animals were equally distributed over the lactation. Days in milk at test-positivity was defined as days from calving to either the date of being tested positive or the end of the lactation (censored observation) for test-negative animals. Risk factors were herd (1 to 8), shedding group (FCNeg, FCLow, FCHigh), and parity (1, 2, and 3). For DIM at FC-positivity as the outcome, antibody status (positive, negative) was included as a risk factor.

First, the descriptive analysis for being test-positive stratified by the risk factors was computed by means of a frequency distribution for ELISA and FC, respectively. The distributions of age and DIM at being test-positive were computed by means of the median and quartiles (Q1 and Q3) for milk and fecal samples, respectively.

A survival analysis using a Cox proportional hazards regression model was used to evaluate the effect of risk factors on age at test-positivity and DIM at test-positivity. In a Cox proportional hazards regression model, no particular form is required for the survival times, and in particular, the baseline hazard function is unspecified. The baseline hazard function is an arbitrary function of time and does not have to be specified explicitly. The Cox proportional hazards regression model is a semiparametric analysis. The single assumption is proportional hazards because the ratio of the hazards for any 2 animals is assumed to be constant over time.

In the survival analyses, the outcome variables were age at test-positivity and DIM at test-positivity. Separate analyses were made for test-positivity, defined by using ELISA on milk samples and FC in fecal samples, respectively. The significance level of the risk factors was computed using the likelihood ratio statistic and a 2 distribution (Collett, 2003). All risk factors were initially included in the model. Interactions between risk factors were not evaluated. Backward elimination of nonsignificant ($P > 0.05$) risk factors was used. The hazards ratio and the 95% confidence intervals for significant risk factors were calculated.

The assumption about proportional hazard functions when using Cox proportional hazards models was evaluated by visual inspection of $\log\{-\log[S(t)]\}$ vs. $\log[S(t)]$ plots, where $S(t)$ is the survival function.

RESULTS

The descriptive statistics with distribution of antibody status and the distribution stratified by shedding group, herd, and parity are shown in Table 1. The distributions of the outcomes, DIM and age, are given in Table 2, divided into antibody-positive and antibody-negative cows. Where the FC was assessed, descriptive statistics of the data are provided in Tables 3 and 4.

Table 1. Distribution of antibody-positive and -negative cows stratified by shedding group, herd and parity

Variable	Level	Antibodies in milk samples			
		Negative		Positive	
		n	(%)	n	(%)
		1,174	(66.1)	602	(33.8)
Shedding Group	FC_{Neg}	1,026	(71.8)	402	(28.2)
	FC_{Low}	86	(42.4)	117	(56.9)
	FC_{High}	62	(42.8)	83	(57.2)
Herd	1	94	(69.1)	42	(30.9)
	2	171	(76.3)	53	(23.7)
	3	167	(61.9)	103	(38.1)
	4	101	(67.3)	49	(32.7)
	5	120	(57.1)	90	(42.9)
	6	319	(65.2)	170	(34.8)
	7	83	(61.9)	51	(38.1)
	8	119	(73.0)	44	(27.0)
Parity	1	407	(66.1)	209	(33.9)
	2	346	(62.2)	210	(37.8)
	≥3	421	(69.7)	183	(30.3)

Table 2. Distribution of age and DIM for antibody-positive and –negative cows.

Variable	Antibodies in milk samples					
	Negative			Positive		
	n	Median	Q_1 to Q_3[1]	n	Median	Q_1 to Q_3
Age (yr)	1,174	4.1	3.3 to 5.3	602	3.9	3.2 to 4.9
DIM	1,174	209	103 to 329	602	217	125 to 301

[1]Q1: 1st quartile 1; Q3: 3rd quartile

Table 3. Distribution of fecal culture positive and negative cows overall and stratified by ELISA-groups, herd and parity

Variable	Level	Fecal culture			
		Negative		Positive	
		n	(%)	n	(%)
		1,428	(80.4)	348	(19.6)
ELISA	Negative	1,026	(87.4)	148	(12.6)
	Positive	402	(66.8)	200	(33.2)
Herd	1	107	(78.7)	29	(21.3)
	2	211	(94.2)	13	(5.8)
	3	249	(92.2)	21	(7.8)
	4	130	(86.7)	20	(13.3)
	5	145	(69.0)	65	(31.0)
	6	374	(76.5)	115	(23.5)
	7	70	(52.2)	64	(47.8)
	8	142	(87.1)	21	(12.9)
Parity	1	404	(72.7)	152	(27.3)
	2	431	(81.3)	99	(18.7)
	≥3	593	(85.9)	97	(14.1)

Table 4. Distribution of age and DIM for fecal culture positive and negative cows.

Variable	Fecal culture					
	Negative			Positive		
	n	Median	Q_1 to Q_3	n	Median	Q_1 to Q_3
Age (yr)	1,428	4.3	3.4 to 5.6	356	3.6	2.9 to 4.7
DIM	1,428	224	98 to 324	348	179	78 to 287

[1]Q1: 1st quartile 1; Q3: 3rd quartile

The resulting multivariable Cox proportional hazards models for age and DIM at detection of antibodies are presented in Table 5. The risk factor "herd" was not significant in either model, whereas "shedding group" was significant in both models. There was no difference between FCHigh and FCLow in either of the models. Cows from the FCLow group were 2.1 times more likely to test positive with ELISA than cows that were never FC-positive, and cows from the FCHigh group were 2.9 times more likely to be ELISA-positive, as indicated by the hazards ratio.

Table 5. Resulting multivariable Cox proportional hazards models of risk factors associated with age and DIM when tested positive based on antibodies in milk samples

Variable	Level	Estimate	SE[1]	HR[1]	95%CI	Overall *P*
Age						
Shedding Group	FC_{Neg}	0	-	1	-	<0.001
	FC_{Low}	0.75	0.11	2.11	1.71-2.59	
	FC_{High}	1.06	0.12	2.88	2.27-3.65	
DIM						
Shedding Group	FC_{Neg}	0	-	1	-	<0.001
	FC_{Low}	0.78	0.11	2.18	1.77-2.68	
	FC_{High}	1.09	0.12	2.98	2.35-3.78	
Parity	1	0	-	1	-	<0.001
	2	0.35	0.09	1.42	1.19-1.70	
	≥3	0.31	0.12	1.37	1.08-1.73	

[1] FC_{High}: culture positive cows with high bacterial counts (> 10 cfu/g) or many repeated FC-positive tests without any non-zero counts of bacteria in-between. FC_{Low}: remaining FC-positive cows; FC_{Neg}: culture negative cows; SE: standard error; HR: hazard ratio; CI: confidence interval

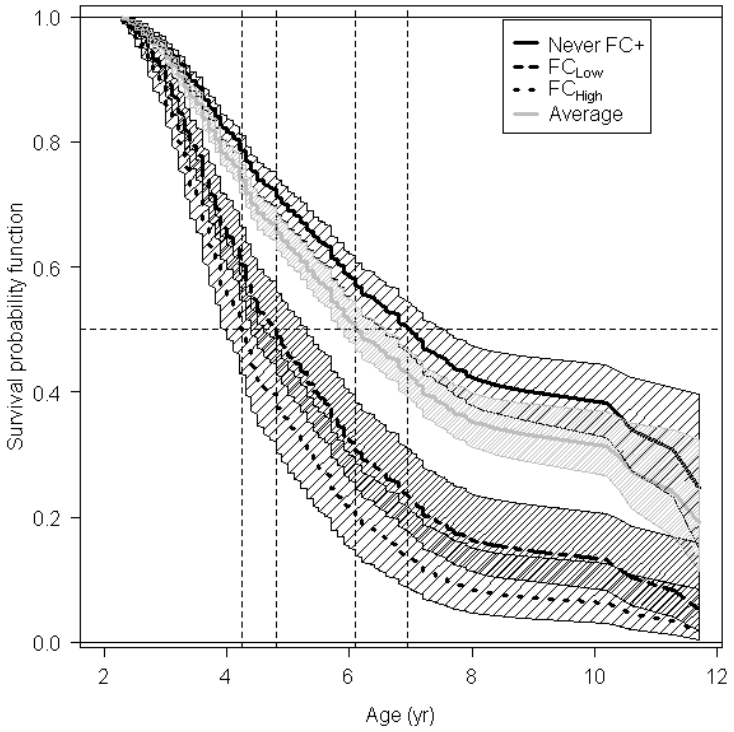


Figure 1. Survival probability plot of age at occurrence of antibodies in cows, average, for cows without detected shedding of *Mycobacterium avium* subsp. *paratuberculosis* (Never FC+), for cows with low level of shedding (FC_{Low}), and for cows with high levels of shedding (FC_{High}). Vertical lines are drawn, for each fecal group, at the age at which 50% of the group had tested positive. For each survivor function, the 95% confidence interval is shown with shaded areas around the graph.

The resulting survival probability functions of age and DIM are shown in Figures 1 and 2, respectively. The median age for being ELISA-positive was 6.1 yr (Q1 to Q3: 4.2 to 10.9). For the FCHigh group, the median age was 4.3 yr (Q1 to Q3: 3.4 to 5.7). At 11.7 yr, 98% (95% confidence interval: 0.91 to 1.0) of the FCHigh cows had become antibody-positive (Figure 1). Thus, 98% for the FCHigh cows developed antibodies. For the FCLow cows, 95% developed antibodies.

In the range of 0 to 350 DIM, the slopes of the survival probability curves for all groups were almost linear (Figure 2), suggesting that the risk of a cow becoming test-positive was equally distributed in this period. After 350 DIM, the slopes were not as steep, indicating that the probability of testing positive decreased late in lactation.

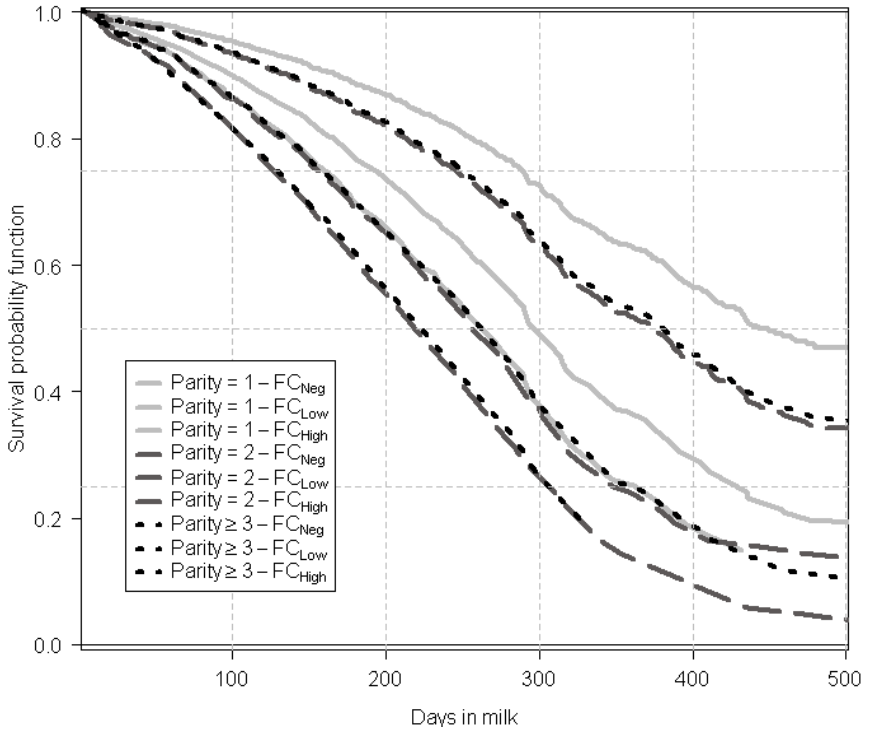


Figure 2. Survival probability plot of DIM at occurrence of antibodies against *Mycobacterium avium* subsp. *paratuberculosis* (Map) in cows, average, by parity, and by shedding group. Black survival functions are cows with high levels of Map (FC_{High}); dark-grey survival functions are cows with low levels of Map (FC_{Low}); and light-grey survival functions are from cows always negative in fecal culture (FC_{Neg}).

Table 6 gives the results for models based on FC. Herd and parity group were not significant. Only the risk factor "antibody-positivity" was significant in both analyses. Antibody-positive cows were 2.3 times more likely to shed detectable amounts of Map with aging, relative to antibody-negative cows. The survival probability functions of age and DIM for FC are shown in Figures 3 and 4, respectively. The median for being FC-positive was at 11.2 yr of age (Q1 to Q3: 5.6 to > 11.3). Of the antibody-positive cows, 50% had become FC-positive at 7.6 yr of age (Q1 to Q3: 4.3 to > 11.3; Figure 3). For ELISA-positive as well as ELISA-negative cows, the probability of being FC-positive was similar from 0 to 300 DIM (Figure 4), whereas few cows tested positive later in lactation.

Table 6. Multivariable Cox proportional hazards models of risk factors associated with age and DIM when tested positive in fecal culture.

Variable	Level	Estimate	SE[1]	HR[1]	95%CI[1]	Overall *P*
Age						
ELISA	Negative	0	-	1	-	<0.001
	Positive	0.83	0.11	2.28	1.84-2.82	
DIM						
ELISA	Negative	0	-	1	-	<0.001
	Positive	1.06	0.11	2.87	2.32-3.55	

[1] SE: standard error; HR: hazard ratio; CI: confidence interval

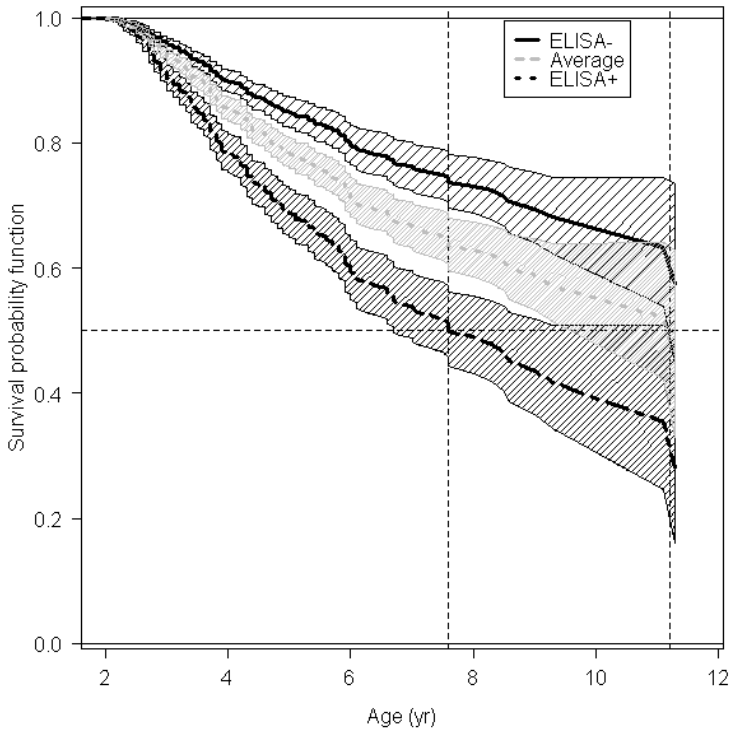


Figure 3. Survival probability plot of age at detection of *Mycobacterium avium* subsp. *paratuberculosis* in cows, average, for cows ELISA-positive, and for ELISA-negative cows. Vertical lines are drawn, for each group, at the age at which 50% of the group had tested positive. For each survivor function, the 95% pointwise confidence interval is shown with shaded areas around the graph.

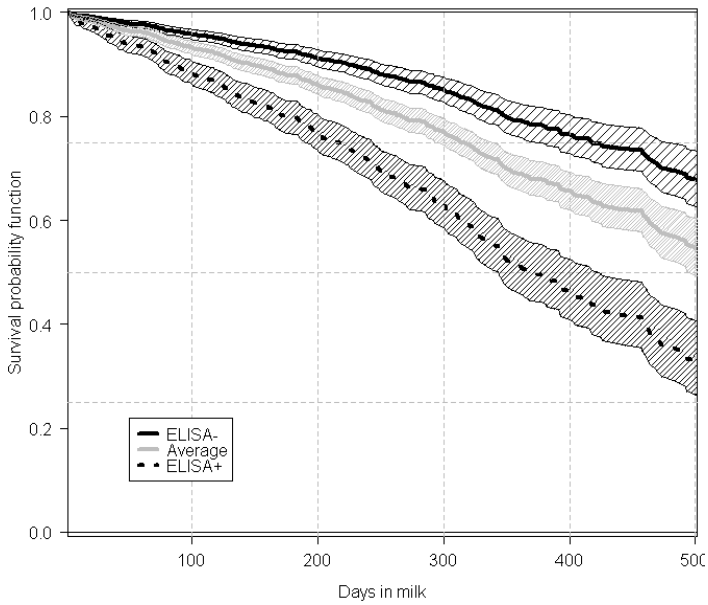


Figure 4. Survival probability plot of DIM at detection of *Mycobacterium avium* subsp. *paratuberculosis* in cows, average and by ELISA group. For each survivor function, the 95% confidence interval is shown with shaded areas around the graph.

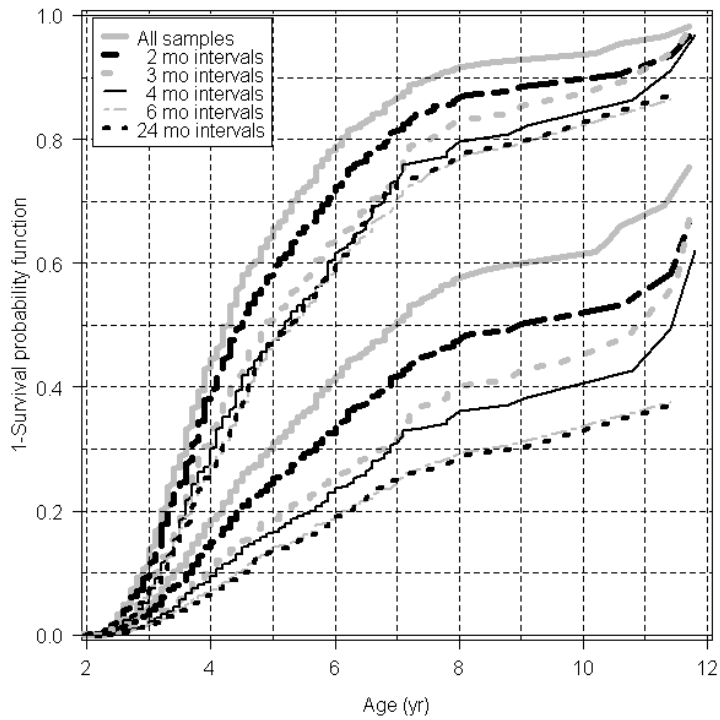


Figure 5. Effect of sampling interval on the probability for testing positive *Mycobacterium avium* subsp. K (Map) (1-survivor probability function) in the milk ELISA with repeated sampling. Sampling frequency varies from 2 to 24 mo (see legends). Top 6 lines are from fecal culture Map positive cows. Lower 6 graphs are from fecal culture Map negative cows.

For repeated ELISA tests, the effect of the interval between sampling on the probability of detecting cows shedding Map is illustrated in Figure 5 by depicting 1 minus the survivorship function. Survivor functions from 4 to 24 mo were almost identical for the FC-positive cows. Hence, only the strategy with all samples included (9 to 10 for a 305-d lactation) and the strategies with 2, 3, 4, 6, and 24 mo between samplings are shown. At yr 3 of age, the drop in the probability of detecting FC-positive cows with ELISA using the 24-mo strategy rather than the "all" strategy was 0.06. At yr 3.5 of age, the drop was 0.13; at yr 4 of age, the drop was 0.18; at yr 5 of age, the drop was 0.18; and at yr 6 of age, the drop was 0.19 (Figure 5), suggesting that a higher sampling frequency is of greater value until yr 4 of age and not later.

Visual inspection of the log{–log[*S(t)*]} vs. log[*S(t)*] plots showed parallel curves when stratified by the different risk factors, indicating that the assumption regarding proportional hazards was fulfilled.

DISCUSSION

A frequent statement made in veterinary medicine textbooks on cattle is that paratuberculosis most frequently occurs in cows between 3 and 5 yr of age (Chiodini et al., 1984; Radostits et al., 2000), which would be expected because most cows are of this age. For example, 65% of Danish dairy cows (n = 671,584) were 3 to 5 yr old on December 31, 2005, based on information from the Danish Central Herd Register. However, time-to-event analyses providing evidence for the statement have not been performed. This statement is partly concordant with the results of the present study: From 2.5 to 4.5 yr of age, the survivorship (i.e., survival probability) function of ELISA-positivity was steepest (for FC-positive cows). For ELISA-positive cows, the survivor-ship function of FC-positivity was steepest from 2.5 to 5.5 yr of age, thereby indicating that the speed of conversion to positivity was highest at this age. Moreover, 40 to 50% of the animals became test-positive after 5 yr of age. Therefore, although many animals do not become test-positive prior to reaching old age, establishing the most cost-effective test strategies requires including the expected test responses.

The immediate conclusion of the present study is that cows are shedding detectable amounts of Map and that they develop antibodies at some point in time (Figure 1). The test sensitivity is higher in the first 300 DIM (Figure 2). One-quarter of the cows with antibodies do not shed detectable amounts of Map (Figure 3). The sensitivity of FC is slightly higher in the beginning of lactation (<300 d; Figure 4). A monthly testing frequency during lactation can increase the probability of detecting Map-shedding cows by ELISA by approximately 18 to 19 percentage points, an increase primarily obtained through high-frequency testing prior to 4 yr of age (Figure 5). This means that with monthly testing of cows <4 yr of age, 42% of cows testing FC-positive sometime during their lifetime would be detected, whereas with twice annual testing, only 24% of the same FC-positive cows would be detected. For older cows, this difference did not exist. For the older cows, the advantage of frequent testing would be earlier detection.

The interpretation of these results may still be a little confusing, primarily because no test has yet been accepted as the gold standard; that is, whether ELISA or FC should be used as the reference has not been determined. Neither test has perfect specificity in infected herds because passive uptake of Map without subsequent infection may lead to false-positive FC, and immune reactions to infections with other mycobacteria may lead to

false-positive ELISA. Thus, neither test should be used as a reference for freedom from infection. The specificity of the ELISA was estimated at 0.99 for cows 2 yr of age and at 0.93 for cows 5 yr of age. The specificity of FC was estimated at between 0.96 and 0.99 during the same age span (Nielsen and Toft, 2006). As a result, a number of cows in the present study were misclassified. With the higher test frequency of ELISA, the serial specificity of ELISA is lower than the serial specificity of FC. Therefore, more cows classified as ELISA-positive in the analysis of the FC response would be misclassified than cows classified FC-positive in the analysis of the ELISA response.

Our findings suggest that cows shedding measurable amounts of bacteria will develop antibodies (Figure 1), but FC cannot detect all cows with antibodies (Figure 3). The utility of the ELISA test in detecting cows shedding high numbers of bacteria has been demonstrated here. Which of the 2 events occurs first is highly variable and is required to determine whether the use of this utility should be operational. A thorough characterization of these aspects is beyond the objectives of this study.

Because antibody production and bacterial shedding occur over the entire age period studied, the diagnostic sensitivity at a given age cannot be very high, which is consistent with the low sensitivity estimates reported by McKenna et al. (2005). Assuming infection in calf-hood, the probability of detecting the shedding of detectable numbers of Map in cows 4 yr of age or older is only 40 to 50% with repeated testing (Figure 1). Age is an important factor in establishing test strategies. The time of testing during lactation appears less important, although ELISA testing later than 350 DIM appears to reduce the sensitivity to some extent (Figure 3). The drop in the probability of testing FC-positive late in lactation is less pronounced (Figure 4). More important is the testing frequency with ELISA, which surprisingly has a primary effect until 4 yr of age. This means that frequent testing of young cows can provide more information than frequent testing of older cows. In young cows, the specificity of the ELISA is highest (Nielsen and Toft, 2006); thus, the drop in specificity would have less effect on the serial specificity than later in life. Monthly sampling of young cows could increase the sensitivity of ELISA and would be advisable, although a cost–benefit analysis to determine the optimal frequency of testing should be performed. The price of an ELISA relative to FC provides additional motivation for the use of this test. In Denmark, the cost of an ELISA, including sampling, is currently DKK 23 (roughly $4.00). The cost of an FC, including sampling, is approximately 185 DKK (roughly $32.00), which is equal to 8 ELISA tests performed at the cost of one FC, and with a huge gain in information. For cows older than 4 yr, less frequent sampling is needed, resulting in additional savings.

This study has some strengths and weaknesses because of its observational nature. The primary strength is that the cows in the study were naturally infected, and selection bias should not be a major factor in assessing the occurrence of antibodies and the shedding of Map. The primary weaknesses are that cows provided a variable number of samples, and followup analyses on slaughterhouse material were not possible, because the cows were sent to multiple slaughter facilities, a factor beyond our control. Postmortem studies on the cows could have provided additional information on their true infection status.

CONCLUSIONS

Almost all cows shedding high numbers of Map also showed antibodies within their lifetime. Cows became test-positive by ELISA and FC during the entire life span, from 2 to 11 yr of age, although the risk of testing positive by ELISA was greatest from 2.5 to 4.5 yr of age.

Cows should preferably be tested in the first 350 DIM. Monthly testing with ELISA could increase the sensitivity substantially for cows <4 yr of age, but no gain was apparent from more frequent testing of older cows, particularly among cows shedding detectable amounts of bacteria.

ACKNOWLEDGMENTS

This study was financially supported by the Danish Cattle Federation (Århus, Denmark).

REFERENCES

Chiodini, R.J., H.J. van Kruiningen, and R.S. Merkal. 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74: 218-262.

Collett, D. 2003. Modelling survival data in medical research, 2.ed, Chapman & Hall, London.

Collins, M.T., S.J. Wells, K.R. Petrini, J.E. Collins, R.D. Schultz, and R.H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12: 685-692.

Hosmer, D.W., and S. Lemeshow, 1999. Applied survival analysis: regression modeling of time to event data. John Wiley and Sons, Inc., N.Y., USA.

Jubb, T., and J. Galvin. 2000. Herd testing to control bovine Johne's disease. Vet. Microbiol. 77: 423-428.

Kalis, C.H.J., J.W. Hesselink, E.W. Russchen, H.W. Barkema, M.T. Collins, I.J.R. Visser. 1999. Factors influencing the isolation of *Mycobacterium avium* subsp. *paratuberculosis* from bovine fecal samples. J. Vet. Diagn. Invest. 11: 345-351.

Lepper, A.W.D., C.R. Wilks, M. Kotiw, J.T. Whitehead, and K.S. Swart. 1989. Sequential bacteriologic observations in relation to cell-mediated and humoral antibody responses of cattle infected with *Mycobacterium paratuberculosis* and maintained on normal or high iron intake. Aust. Vet. J. 66: 50-55.

McKenna, S.L.B., G.P. Keefe, H.W. Barkema, and D.C. Sockett. 2005. Evaluation of three ELISAs for *Mycobacterium avium* subsp. *paratuberculosis* using tissue and fecal culture as comparison standards. Vet. Microbiol. 110: 105-111.

Nielsen, S.S. 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subspecies *paratuberculosis* in dairy cattle. J. Vet. Med. B 49: 384-387.

Nielsen, S.S., and N. Toft. 2006. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. J. Dairy Sci. 89: 569-579.

Nielsen, S.S., B. Kolmos, and A.B. Christoffersen. 2004. Comparison of contamination and growth of *Mycobacterium avium* subsp. *paratuberculosis* on two different media. J. Appl. Microbiol. 96: 149-153.

Nielsen, S.S., C. Enevoldsen, and Y.T. Gröhn. 2002. The *Mycobacterium avium* subsp. *paratuberculosis* ELISA response by parity and stage of lactation. Prev. Vet. Med. 54: 1-10.

Radostits, O.M., C.C. Gay, D.C. Blood, and K.W. Hinchcliff. 2000. Veterinary Medicine, 9th Edition. W.B. Saunders Company, London, UK.

Stabel, J.R. 2000. Transistions in immune responses to *Mycobacterium paratuberculosis*. Vet. Microbiol. 77: 465-473.

Sweeney, R.W., R.H. Whitlock, C.L. Buckley, and P.A. Spencer. 1995. Evaluation of a commercial enzyme-linked immunosorbent assay for the diagnosis of paratuberculosis in dairy cattle. J. Vet. Diagn. Invest. 7: 488-493.

Toft, N., S.S. Nielsen, and E. Jørgensen. 2005. Continuous-data diagnostic tests for paratuberculosis as a multistage disease. J. Dairy Sci. 88: 3923-3931.

Van Schaik, G. C.R. Rossiter, S.M. Stehman, S.J. Shin, and Y.H. Schukken. 2003. Longitudinal study to investigate variation in results of repeated ELISA and culture of fecal samples for *Mycobacterium avium* subsp. *paratuberculosis*. Am. J. Vet. Res. 64: 479-484.

Paper III

Continuous-data diagnostic tests for paratuberculosis as a multi-stage disease

Toft N, Nielsen SS, Jørgensen E

Published in
Journal of Dairy Science 2005, 88, 3923–3931

Continuous-data diagnostic tests for paratuberculosis as a multi-stage disease

N. Toft,[1] S. S. Nielsen,[1] E. Jørgensen[2]

[1]Department of Large Animal Sciences, The Royal Veterinary and Agricultural University, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark
[2]Biometry Research Unit, Danish Institute of Agricultural Sciences, P.O. Box 50, DK-8830, Tjele, Denmark

ABSTRACT

We devised a general method for interpretation of multistage diseases using continuous-data diagnostic tests. As an example, we used paratuberculosis as a multistage infection with 2 stages of infection as well as a noninfected state. Using data from a Danish research project, a fecal culture testing scheme was linked to an indirect ELISA and adjusted for covariates (parity, age at first calving, and days in milk). We used the log-transformed optical densities in a Bayesian network to obtain the probabilities for each of the 3 infection stages for a given optical density (adjusted for covariates). The strength of this approach was that the uncertainty associated with a test was imposed directly on the individual test result rather than aggregated into the population-based measures of test properties (i.e., sensitivity and specificity).

INTRODUCTION

Traditionally, the interpretation of a diagnostic test is a positive or a negative test result and the uncertainty associated with the test usually is measured by the sensitivity and specificity (i.e., the proportions of truly diseased and truly nondiseased identified correctly by the test). In this approach, however, limiting assumptions include 1) a dichotomous disease definition; 2) a threshold (cut-off) clearly dividing positive and negative test results; 3) all positive (and negative) test results are equally positive (or negative); and 4) the animals' test response is not affected by relevant covariates.

Often the test is carried out merely to determine whether a specific condition is present to initiate a suitable intervention. For this purpose, dichotomizing the disease definition and test result is adequate. It might be worthwhile, however, to improve this approach when the disease or condition and the associated threshold is ambiguous, thus allowing tests to be used in a wider range of settings.

Consider, as an example, paratuberculosis: a chronic, slowly developing infection in cattle and other ruminants caused by *Mycobacterium avium* ssp. *paratuberculosis* (Map; Chiodini et al., 1984). The chronicity of infection makes simple definitions of disease difficult. Furthermore, the sensitivity of diagnostic tests varies depending on the stage of the disease (Nielsen et al., 2002c). A frequently adopted intervention for infected cows is culling as opposed to doing nothing. However, some infected cows never develop "clinical" paratuberculosis with diarrhea and concomitant emaciation. Some managers only cull those cows that experience clinical disease, whereas other managers would like to detect and cull subclinical animals that might transmit Map to herd mates, are less productive, or both. Thus, a general framework to devise an optimal test-and-cull policy for paratuberculosis would benefit from a test that allows for multiple classifications of infection. Nielsen et al. (2002b)

suggested 3 stages of Map infection: noninfected cows, infected cows with predominating cell-mediated immune responses, and infected cows with predominating humoral immune responses. The infected cows with predominating cell-mediated immune responses are assumed to have reduced antibody titers during the primary cell-mediated immune responses. During humoral immune responses, antibody titers are expected to be elevated. Hence, the 3 "infection groups" defined above may be assumed to correspond to 3 "immuno-groups": noninfected (having no antibodies); infected, with reduced antibody titers; and infected, with elevated antibody titers. Validating such immuno-groups requires repeated testing using fecal culture (FC), which is time consuming and expensive. Fecal culture generally takes 12 wk and sampling requires extra work compared with a milk-based indirect ELISA. However, given that a link between the immuno-groups and the infection groups exists, antibody testing could provide a tool for inference about the amount of bacterial shedding, and hence, be used as decision support to livestock producers rather than using the more cumbersome and expensive FC method.

Other influences (such as parity and stage of lactation) on optical densities (OD) of the milk ELISA have been demonstrated previously (Nielsen et al., 2002c). Statistical models should include these covariates to improve the interpretation of the test result. Furthermore, such models should be easily adapted to repeated measures of test results on individual cows because this might be an element of a future testing regimen. Standard statistical methods such as mixed linear normal models (e.g., PROC MIXED, v. 8.2; SAS Inst., Inc., Cary, NC) are well established for inference from such models. Rather than OD given infection status, for diagnostic inference the conditional distribution of the infection status given the observed OD and relevant covariates of the cow is required. This becomes possible with comparatively little effort using probabilistic expert systems, such as Bayesian networks (Cowell et al., 1999).

The objective of the current study was to demonstrate how continuous-test data (and covariates) can be used for diagnostic testing when diseases are allowed to have multiple stages, as exemplified by an ELISA used for classifying lactating dairy cows into the 3 infection stages of paratuberculosis as defined above.

MATERIALS AND METHODS

Sampling Frame

For the current study, farms from a region in Southern Jutland, Denmark, where the Danish dairy industry initiated a project on infectious diseases in 1998 (Andersen et al., 2000) were used. The region was defined by 4 postal codes and contained approximately 260 dairy herds in 1998. Among these herds, 110 participated in the project. Specifically, results from 5 herds in which all cows had tested negative for paratuberculosis in 3 consecutive rounds of FC, based on samples collected 1 yr apart, and 8 herds in which several cows from each herd were shedding Map were used. These latter 8 herds were visited every third month between December 1999 and December 2002. Fecal samples were collected from all cows present at the time of the visit. Milk samples were collected 11 times/yr from each herd through the routine milk production scheme.

Test Methods and Classification Scheme

Fecal samples were cultured for 12 wk and classified as either negative or positive with varying degrees of bacterial growth. Positive cultures were confirmed by PCR detecting

IS*900*. The test was described in detail in Nielsen et al. (2004). Based on these classifications, the cows were trichotomized according to their Map infection status: 1) cows in 5 herds that never had any culture-positive cows were assumed free of paratuberculosis and all tested cows from these herds were classified as fecal-culture negative (FC_{neg}); 2) from the 8 herds with known problems, FC-positive cows with some negative cultures and whose positive cultures always had only few counts of bacteria (< 10 cfu/g of feces) were defined as fecal culture low (FC_{low}); and 3) from the 8 herds with known problems, FC-positive cows with elevated bacterial counts (10 cfu/g of feces) or many repeated FC-positive tests (with any nonzero counts of bacteria) were defined as fecal culture high (FC_{high}). The latter group of cows can be perceived as those failing to control the Map infection.

Culture-negative cows from the 8 herds with known problems were designated "fecal-culture unidentified" and excluded from further analysis. The poor sensitivity of the FC prohibited a reasonable diagnosis of these cows. Based on this classification, 3983 results from 737 cows were included in the study.

The milk samples were tested in an ELISA for antibody based on *Mycobacterium paratuberculosis* strain 18 (an *M. avium* ssp. *avium* strain) from Allied Monitor (Fayette, MO). This test was described in detail previously (Nielsen et al., 2001, 2002c). The test results of the samples were classified based on their OD value from the ELISA reader. The OD was corrected for interplate variation by subtracting the OD of a negative control tested in each ELISA plate. Information on DIM, current parity, and the age at first calving of the tested cows was extracted from the Danish Cattle Database.

Statistical Models

Initial analyses showed that the variance of the OD values increased with increasing mean value. Thus, the corrected OD were log-transformed to stabilize their variance. Initially, a mixed linear normal model was formulated using random effects to take into account repeated measurements. Subsequent tests of model assumptions showed that variance homogeneity could not be assumed, and a more detailed variance model had to be formulated. The detailed model specification follows.

Paratuberculosis is a chronic infection, and infection is often assumed to occur during calfhood. Thus, age was included in the model by using the proxies parity, stage of lactation (Nielsen et al., 2002a), and age at first calving. Because the association between log OD and age did not seem to be linear, however, DIM and age at first calving were coded as ordinal variables. Days in milk was divided into 0 to 1, 2 to 11, 12 to 27, 28 to 40, and > 40 wk after calving. Age at first calving was dichotomized into 28 and > 28 mo. Parity was modeled as first, second, and third or greater parity. Although these variables were expected to influence OD for FC_{low} and FC_{high} cows, the same effect was not expected for FC_{neg} cows. Therefore, interaction between FC classification and the age covariates was expected. Furthermore, variance heterogeneity was expected between FC types, necessitating residual variances being allowed to vary. Variation among cows within herds was included as a random effect of cows nested within herds.

The initial model contained all main effects and their first-order interactions and assumed homogeneous variance. Model selection was performed by adding variance heterogeneity among subgroups when significant (tested by the likelihood-ratio test, = 0.05) and sequentially removing nonsignificant (using a 2-tailed test with = 0.05) fixed effects (by using

SAS type III test using Satterthwaite's approximation for calculation of the degrees of freedom for the test). This procedure produced the following model for the log-transformed OD [log(OD)]:

$$\log(\mathrm{OD})_{himn} = \mu + \mathrm{FC_i} + \mathrm{P_{J_{himn}}} + \mathrm{DIM_{K_{himn}}} + \mathrm{AC_{L_{him}}} + (\mathrm{FC}\times\mathrm{P})_{iJ_{himn}} + (\mathrm{FC}\times\mathrm{DIM})_{iK_{himn}} + (P\times AC)_{J_{himn}L_{him}} + [\mathrm{COW}(\mathrm{FC}\times\mathrm{HERD})]_{him} + \varepsilon_{himn} \quad (1)$$

In this model,

$\log(\mathrm{OD})_{himn}$ is the log-transformed corrected OD of the *n*th recording of the *m*th cow in herd *h* within the *i*th FC type.

$\mathrm{FC_i}$ is the systematic effect of the *i*th FC type, i = 1, 2, 3;

$\mathrm{P_{J_{himn}}}$ is the systematic effect of parity, J = 1, 2, 3, with J = 3 indicating cows with parity >2;

$\mathrm{DIM_{K_{himn}}}$ is the systematic effect of the Kth days in milk group, K = 1, ... , 5;

$\mathrm{AC_{L_{him}}}$ is the systematic effect of the age at first calving, L = 1, 2 (subscripts L_{him} indicate that age at first calving is constant for a cow);

$[\mathrm{COW}(\mathrm{FC}\times\mathrm{HERD})]_{him} \sim N(0, \sigma_C^2)$ is the random effect of cow within herd and FC type; and

$\varepsilon_{himn} \sim N(0, \sigma_\varepsilon^2[\exp(\mathbf{U}\delta)]_{iJ_{himn}})$, i.e. residuals are identically Normal distributed within each combination of FC type and parity with σ_ε^2 as the common intercept term of the residual variance, $\mathbf{U}$ as the design matrix reflecting the combination of FC type and parity and δ as an eight-dimensional vector of estimated parameters.

$$\mathbf{U} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Thus, the expression $\sigma_\varepsilon^2[\exp(\mathbf{U}\delta)]$ gives a 9-dimensional vector: the first 3 elements give the variance for parity 1, 2 and ≥3 for FC_{neg} cows, the next 3 elements give the same variances for FC_{low} cows, and the last 3 give variances for FC_{high} cows. Hence, given estimates of σ_ε^2 and δ the variance for FC_{neg} and first-parity cows can be calculated as $\sigma_\varepsilon^2 \exp(\delta_1 + \delta_3 + \delta_5)$, the variance for FC_{low}, first parity as $\sigma_\varepsilon^2 \exp(\delta_2 + \delta_3 + \delta_7)$, etc.

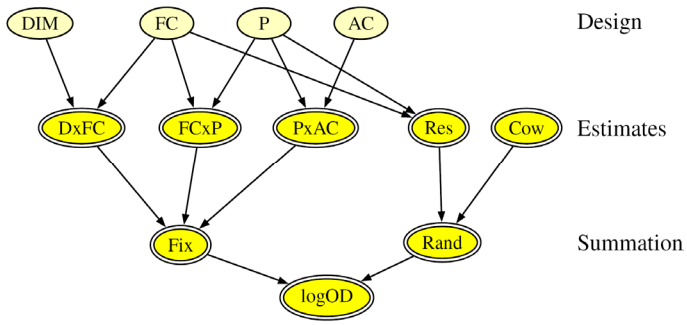


Figure 1. Bayesian network representation of the statistical model (Eq. 1): discrete nodes are represented as single bordered ellipses, whereas the continuous Gaussian nodes are represented as double-bordered ellipses. The design effects include DIM, fecal-culture type (FC), parity (P) and age at first calving (AC). Estimates of these effects are represented in D × FC, FC × P and P × AC for the interaction terms of the fixed effects, and herd effect (cow) and residual variation (res). The nodes Fix and Rand are summation nodes and logOD is the log transformed optical density (OD) of the milk ELISA

Bayesian Network Model

Parameter estimates from the statistical model from equation 1 were used directly to form a Bayesian network (also known as a probabilistic network; Cowell et al., 1999). The qualitative part of this Bayesian network is shown in Figure 1. Bayesian networks may be constructed from expert knowledge concerning the domain. However, a network that corresponded closely to the statistical model with intermediate nodes to model the 2-way interactions was preferred. A special kind of Bayesian network, a Continuous Gaussian graph, where nodes are allowed to be either discrete (ellipses with solid borders) or continuous (Gaussian; ellipse with double lined border) nodes were used (Cowell et al., 1999). The arrows from one node (the parent) to another (the child) indicate a description of the conditional distribution of the values in the node given the values of its parents. For example, in Figure 1, DIM and FC are "parents" of the D x FC node. The discrete nodes were used for modeling the categorical design variables from the model (e.g., whether parity was at level 1, 2, or 3). The first level of continuous nodes was used to represent the parameter estimates and the further levels to represent the subsequent summation of the random and fixed effects. The intermediate summation might have been omitted but facilitates the modeling when repeated measurements on the same cow are modeled. Each continuous node has an associated table with the means and variances for the Normal distribution for each possible configuration of the parent nodes, thus utilizing the standard errors as well as the estimated parameters. For example, the line in the table describing the distribution of the continuous node (FC x P) with fecal type FC_{neg} and first parity would contain the estimate of the parameter $(P \times FC)_{(1,FCneg)}$ = – 0.08 from the model in equation 1 as mean and the square of the corresponding SE (= 0.03) as variance (Table 1). The res node (representing the residual) would have a mean of 0 and the square of the SD [= 0.26 for fecal type FC_{neg} and parity 1 (Table 2)] as variance. In Figure 1, the intercept and the main effects are pooled within the interactions.

Table 1. Estimates of the fixed effects of the model describing the log-transformed corrected optical density (OD), using fecal culture types (FC), parity (P), DIM, and age at first calving (AC). The interaction terms have absorbed the main effects to ease the transformation from the model to a Bayesian network.

Interaction	Variable	Level	Estimate	SE	*P* value
DIM × FC	Weeks				
	0-1	FCneg	−0.60	0.05	<.0001
	2-11		−0.67	0.03	<.0001
	12-27		−0.69	0.03	<.0001
	28-40		−0.66	0.03	<.0001
	>40		−0.62	0.03	<.0001
	0-1	FClow	−0.25	0.07	0.0003
	2-11		−0.31	0.04	<.0001
	12-27		−0.25	0.04	<.0001
	28-40		−0.15	0.04	0.0005
	>40		−0.04	0.05	0.3818
	0-1	FChigh	0.12	0.07	0.0801
	2-11		0.03	0.04	0.4914
	12-27		0.13	0.04	0.0033
	28-40		0.25	0.05	<.0001
	>40		0.40	0.05	<.0001
P × FC	Parity				
	1	FCneg	−0.08	0.03	0.0114
	2		−0.00	0.03	0.7335
	≥3		0	.	.
	1	FClow	−0.45	0.05	<.0001
	2		−0.17	0.04	<.0001
	≥3		0	.	.
	1	FChigh	−0.57	0.04	<.0001
	2		−0.28	0.04	<.0001
	≥3		0	.	.
P × AC	Parity				
	1	≤ 28 mo	−0.05	0.03	0.0659
	1	> 28 mo	0	.	.
	2	≤ 28 mo	0.02	0.03	0.5488
	2	> 28 mo	0	.	.
	≥3	≤ 28 mo	0.08	0.03	0.0146
	≥3	> 28 mo	0	.	.

Table 2. The estimated residual standard deviations for different combinations of fecal culture (FC) type and parity

	Fecal culture type		
Parity	FC_{neg}	FC_{low}	FC_{high}
1	0.26	0.40	0.40
2	0.18	0.34	0.44
≥3	0.25	0.32	0.38

In each node of a Bayesian network the observed value of a variable (evidence; i.e., the parity of a cow or the log OD value), may be entered. Using probabilistic calculus, the network reevaluates the probability distribution of the nodes in the network based on the evidence. This reevaluation is termed propagation of evidence. As an example, given evidence on every design variable, the network calculates the predicted log OD value with corresponding standard error. If no evidence is available, the network uses the prior distribution for the calculation. Prior distributions are often chosen to be noninformative (for example, equal probabilities for each level of the design variable). If evidence in the parity node were omitted, the calculated log OD would essentially correspond to least squares means in SAS terminology. As with least squares means, it is typically better to use a weighting that corresponds to the distribution over parity in the population rather than equal weighting. In the Bayesian network, this corresponds to assigning a prior distribution to the parity node that corresponds to population distribution. When evidence (information) about a variable is entered into a Bayesian network, the prior distribution for that variable is disregarded. The strength of the Bayesian network is that it is possible to make deductions against the arrows. Upon entering evidence about the log OD, a revised probability distribution of the nodes without evidence is obtained. In the case of diagnostic information, the probability of a cow belonging to the FC_{neg}, FC_{low}, or FC_{high} category based on the measured log OD value is the primary concern. Thus, using a Bayesian network allows inference about the probability (Pr)(FC_{neg}), Pr(FC_{low}), and Pr(FC_{high}) for a given log OD, using age at first calving, parity, and stage of lactation for a cow as age covariates, while accounting for the uncertainty associated with the parameter estimates. The inference made by propagations in the Bayesian network is essentially an application of Bayes formula (Cowell et al., 1999). Thus, it was possible to specify the "natural" model (i.e., FC type explaining log OD) in a traditional setting of mixed models ANOVA using a frequentist approach with well-established methods for model selection, and subsequently obtain the relevant information (log OD explaining FC type).

Because age at first calving, parity, and stage of lactation are known, evidence will overrule the prior, and uniform distributions may be used. The prior distribution on the proportion of FC_{neg}, FC_{low}, and FC_{high} cows should be given more attention, ideally reflecting the distribution in the population where the Bayesian network will be applied. Distribution in data, however, is not necessarily representative; a herd-specific prior might even be required. Hence, for the illustrative purposes of this study, the prior distribution of FC type is also assumed uniform.

RESULTS

In Figure 2, histograms for the log OD of each FC type and the empirical distribution of FC_{neg}, FC_{low}, and FC_{high} with increasing log OD are given. The probability that a cow with a log OD of –1 is FC_{neg} is nearly 40% [the vertical distance between the lower horizontal axis and the first (dotted) line], whereas Pr(FC_{low}) for log OD above 0 is near 0%. The probability of FC_{low} is almost constant, but Pr(FC_{high}) increases with increasing log OD. The empirical distribution is shown using stacked probability plot to emphasize that Pr(FC_{neg}) + Pr(FC_{low}) + Pr(FC_{high}) = 1 for any given log OD.

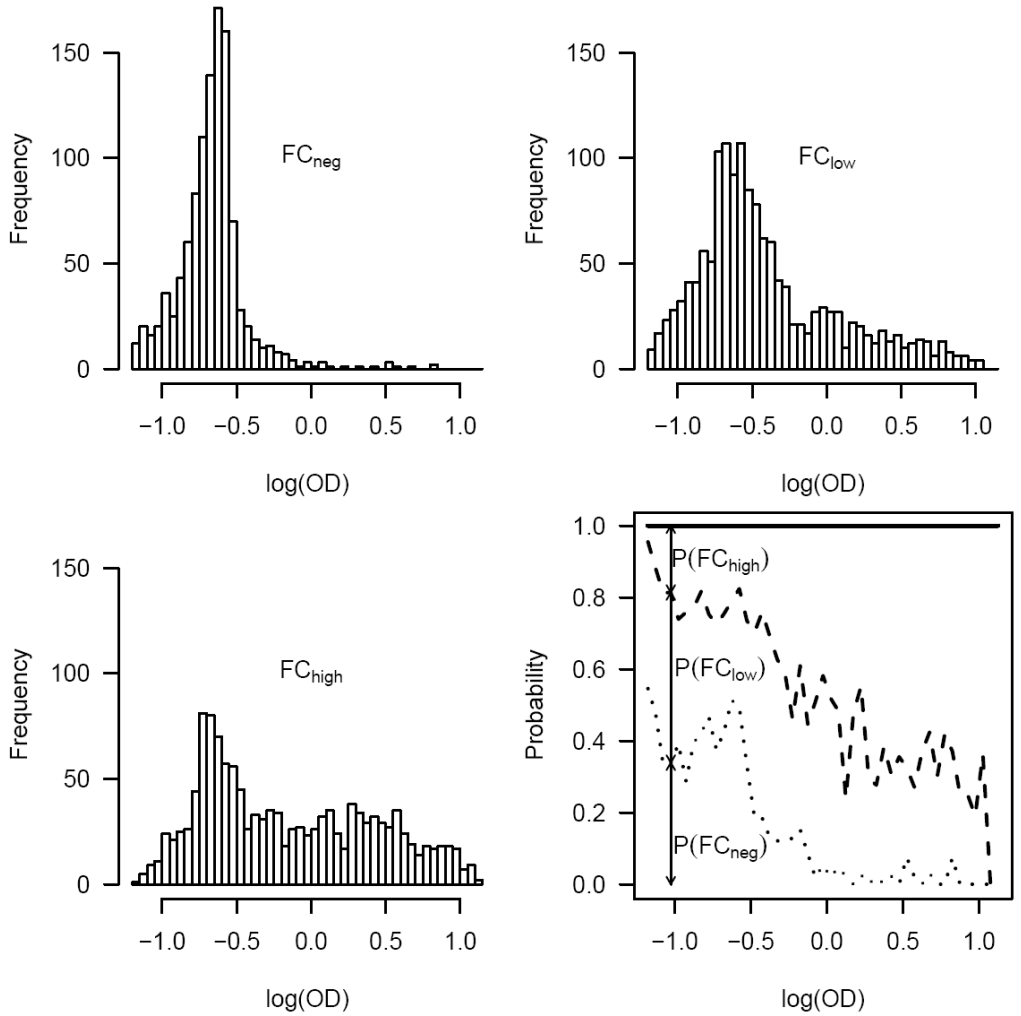


Figure 2. Histograms of the log-transformed corrected optical density (OD) of the ELISA by fecal culture (FC) types, and a plot of the empirical distribution of the 3 FC types as a function of the log-transformed corrected OD. Pr(FC_{neg}) of a given log(OD) is the distance between the lower horizontal axis and the dotted line, the distance between the dotted line and the dashed line is Pr(FC_{low}), and the distance between the dashed line and the solid upper horizontal line is Pr(FC_{high}).

Table 1 gives the parameter estimates and SE used in the quantitative part of the Bayesian network (Figure 1). The parameterization of the Bayesian network was done without the main fixed effects. Hence, these were excluded (and consequently absorbed by the interaction terms) when estimating the solutions to the fixed effects. Although significant effects of the statistical model in equation 1 and the estimated parameters (Table 1) were only used for constructing the Bayesian network (Figure 1), it seems relevant to review briefly whether the results found here are in accordance with results found in previous studies. Using SAS Type III test with Satterthwaite's approximation to calculate the degrees of freedom for the tests, significant effects of the interaction between the DIM and fecal-culture type ($P < 0.001$), between fecal culture-type and parity ($P < 0.001$), and between parity and

age at first calving (P < 0.002) were found. The mean log OD is increasing from FC_{neg} to FC_{low} and FC_{high}; all 3 means were significant when tested using differences in least squares means (results not shown). This supports the hypothesis that the 3 classifications (FC_{neg}, FC_{low}, and FC_{high}) represent 3 different stages of the infection. In the presentation of the estimates, as well as in the Bayesian network, the possible correlation between the estimates has been ignored.

The estimate of the herd effect was σ_C^2 = 0.0712 (SE = 0.0053; *P* < .0001). Given the estimates of the common intercept term of the residual variance (σ_ε^2 = 0.14) and the vector δ = (−0.84,−0.32, 0.12, 0.30, 0.02,−0.96, 0.30,−0.18) the residual variance for the individual combinations of FC type and parity was calculated as:

$\sigma_\varepsilon^2\left[\exp(\mathbf{U}\delta)\right]$ = (0.0698, 0.0316, 0.0610, 0.1565, 0.1161, 0.1030, 0.1594, 0.1917, 0.1414).

The square roots of these estimates are given in Table 2. The estimated SD showed more variation in the log OD for FC_{high} than FC_{low} and FC_{neg} cows within each parity group. The main difference in variance, however, was due to differences between FC_{neg} and the other 2 classes (as expected). Variation between measurements of cows greater than second parity seemed to be less than between first-parity cows.

Using the estimates of Table 1, the random herd effect, and the calculated residual SD (Table 2), the quantitative part of the Bayesian network shown in Figure 1 was constructed. Using this network, the distribution of FC_{neg}, FC_{low}, and FC_{high} for any given log OD, parity, age at first calving, and DIM could be estimated. As each combination gives a different probability (and log OD is a continuous variable), the results of these propagations are best presented as graphs (Figure 3) using stacked plots to emphasize that $Pr(FC_{neg}) + Pr(FC_{low}) + Pr(FC_{high}) = 1$ for any given combination of log OD and age covariates. In Figure 3, the probabilities of FC type for a given log OD are given for 9 different combinations of age covariates (the remaining combinations showed similar results, but were omitted to save space). For a given log OD, $Pr(FC_{neg})$, $Pr(FC_{low})$, and $Pr(FC_{high})$ differ substantially among parities. This is perhaps best seen when comparing small log OD and the associated probability $Pr(FC_{neg})$ cow. Although the result for first-parity cows seemed inconclusive, the result for cows of third or greater parity indicated that small log OD are more likely to be the result of an FC_{neg} cow.

DISCUSSION

The statistical analysis confirmed the results from Nielsen et al. (2002b). The OD is affected by age, with an increase in log OD with age in general, but also within lactations. Allowing local effects of residual variances between FC types and parities confirmed that the variance of the ODs of noninfected cows (FC_{neg}) is smaller than for infected cows (FC_{low} and FC_{high}). The SD of the OD of FC_{neg} cows equals 40 to 80% of the SD of FC_{low} and FC_{high} cows when comparing FC types within parities.

Using a Bayesian network representation of the statistical model allowed the uncertainty associated with the estimates of the statistical model to be represented as variance on the Gaussian nodes. Hence, the estimates of probabilities provided by the Bayesian network are estimates based on the estimated parameters and their associated uncertainty. This implies that the resulting probabilities gives a more realistic picture of the properties of the ELISA as a tool for classification of cows with respect to their bacterial shedding status than a traditional approach in which the misclassification is presented in terms of sensitivity and

specificity of a test. Although the uncertainty associated with these estimates is often acknowledged, it is usually not used in the further interpretation of the results.

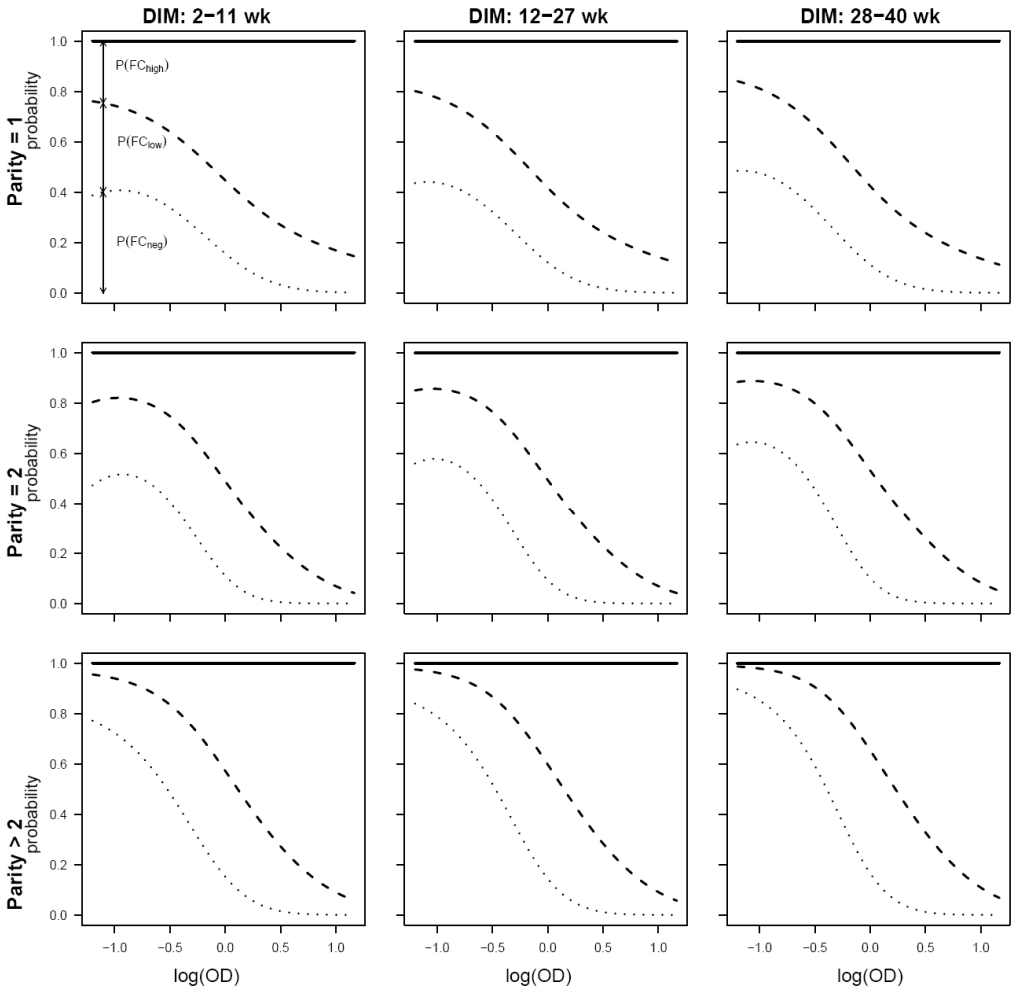


Figure 3. Pr(FC_{neg}), Pr(FC_{low}) and Pr(FC_{high}) as a function of the log-transformed corrected optical density (OD). Age at first calving is ≤ 28 mo for each plot, the vertical distance from the lower horizontal axis to the dotted line is Pr[FC_{neg}| log(OD)], the vertical distance between the dotted and the dashed lines is Pr[FC_{low}| log(OD)], and the vertical distance between the dashed line and the upper horizontal line is Pr[FC_{high}| log(OD)].

When using a trichotomized infection status rather than the usual dichotomization, the concepts of sensitivity and specificity lose their intuitive interpretation. Similarly, so does the use of likelihood ratios such as in Collins (2002), although, if the objective only was to handle continuous tests, they could be used. The Bayesian network and graphical representation of probabilities, however, could just as well be adapted when the traditional dichotomous disease classification is used (or generalized into more than 3 disease stages). Using Bayesian networks for decision support does not require construction of stacked probability

plots, which are only used here for illustrative purposes. The statistical model showed that covariates influence the relationship between FC type and OD. Such covariates should be taken into account when interpreting the diagnostic test. This is further discussed in Nielsen and Toft (2002) for a traditional dichotomous disease definition. Consequently, computer-based methods are needed when interpreting the OD. This problem is not unique to the current study and rather than trying to further justify the use of covariates for test interpretation, it is probably more appropriate to challenge the justification in ignoring these known covariates. Visual comparison of the empirical distribution in Figure 2 and the individual distributions of Figure 3 seemingly shows an effect (gain) by including covariates. The plots in Figure 3 differ visually between different configurations of parity and DIM category. Whether these fluctuations are large enough to alter a culling decision based on an OD is another question that will be left for future studies.

The diagnostic properties of tests for Map are rather poor when used for just 2 stages of infection, and adding a third infection stage is unlikely to improve this. Throwing away information by dichotomizing or trichotomizing the test interpretation will further decrease the value of the test. To illustrate this, consider the situation shown in the lower left plot of Figure 3 and assume that 2 cut-offs were introduced so that log OD –0.5 was interpreted as FC_{neg}, –0.5 < log OD 0.5 as FC_{low}, and 0.5 < log OD as FC_{high}. Using this approach, it would be possible to estimate parameters describing the probability of correctly classifying into each of the 3 categories. However, each log OD below –0.5 would have the same probability of being correctly (or wrongly) classified, whereas using the log OD directly would give more confidence of a cow being FC_{neg} when the log OD equals –1.0 compared with –0.55.

Eventually, it comes down to whether an interpretation of a cow being 70% FC_{neg}, 20% FC_{low}, and 10% FC_{high} is manageable. The alternative is to classify the cow as 1 of the FC types and accept that misclassification occurs. When used in a computer-based decision support system, there is no need to know the exact state of the cow with respect to infection with Map. Thus, there seems to be more benefit in representing the uncertainty associated with a given test result directly than through the traditional concept of test properties (such as sensitivity and specificity, which cannot be used here). Still, for a given set of conditions there is an optimal cutoff value for the OD in which all test results above that OD results in an action (such as culling). The point is that a continuous test (such as OD) combined with additional information allows this cutoff value to be assessed for different combinations of age at calving, parity, DIM, and with respect to a specific purpose of the test (i.e., control vs. eradication).

Testing and culling dairy cows according to their Map infection status should ideally be combined with the traits traditionally used to assess the value of a dairy cow in replacement models (e.g., milk yield, reproductive problems, and age); this is reviewed by Kristensen (1994). In Houben et al. (1994), a model that includes health traits such as mastitis were used for optimal replacement of dairy cows. Recently, Gröhn et al. (2003) developed a model that considered several different diseases when optimizing the replacement of dairy cows. These replacement models all assume that relevant information can be observed with certainty. This assumption, however, cannot be justified when addressing paratuberculosis. The true bacterial shedding state (or FC type) of the cow is generally not known and requires extensive and repeated testing (as described earlier in this paper) due to intermittent shedding and low diagnostic sensitivity in the early stages of infection. Fecal sampling and culturing is also more expensive and time consuming compared with milk sampling and

testing using ELISA. Milk samples are already retrieved within the Danish milk-recording scheme for 88% of the Danish dairy herds. Thus, it might be tempting to apply the OD directly in the modeling without using the FC types. However, when modeling the effects of paratuberculosis on the risk of transmitting the infection to other animals, the FC types make more biological sense than the OD. Furthermore, methods for handling the partial observability within a traditional sequential decision-support framework with an infinite time horizon are being developed (Nilsson and Kristensen, 2002).

The present study uses a classification of cows into 3 different categories with respect to paratuberculosis infection status. The procedure used for classification as well as the interpretation of the 3 groups can be questioned. The low sensitivity of the FC might introduce potential bias. However, the primary purpose of this study was to demonstrate that the use of dichotomized disease status to interpret a test result and the use of sensitivity and specificity to evaluate the assay method has shortcomings that make it worthwhile to consider other ways of interpreting and representing the uncertainty of diagnostic tests. The framework presented herein can easily be extended to allow for repeated tests, multiple diseases, or both. If the disease status cannot be determined by elaborate test schemes, then the statistical analysis needs to be carried out as a latent class analysis, such as the mixture model used in Nielsen et al. (2003).

ACKNOWLEDGMENTS

The Danish Dairy Board and the Kongeåproject are thanked for their financial support of the present study.

REFERENCES

Andersen, H., K. Aagaard, F. Skjøth, E. Rattenborg, and C. Enevoldsen. 2000. Integration of research, development, health promotion, and milk quality assurance in the Danish dairy industry. Pages 258-260 in Proceedings of the 9th symposium of the International Society of Veterinary Epidemiology and Economics, Breckenridge, Colorado, Aug 6-11.

Chiodini, R. J., H. J. van Kruiningen, and R. S. Merkal. 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74:218–262.

Collins, M. T. 2002. Interpretation of a commercial bovine paratuberculosis enzyme-linked immunosorbent assay by using likelihood ratios. Clin. Diagn. Lab. Immun. 9:1367–1371.

Cowell, R. G., A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. 1999. Probabilistic Networks and Expert Systems. (Statistics for Engineering and Information Science). Springer Verlag, New York, USA.

Gröhn, Y. T., P. J. Rajala-Schultz, H. G. Allore, M. A. DeLorenzo, J. A. Hertl, and J. A. Galligan. 2003. Optimizing replacement of dairy cows: modeling the effects of diseases. Prev. Vet. Med. 61:27–43.

Houben, E. H. P., R. B. M. Huirne, A. A.Dijkhuizen, and A. R., Kristensen. 1994. Optimal replacement of mastitic cows determined by a hierarchic Markov process. J. Dairy Sci. 77:2975–2993.

Kristensen, A. R. 1994. A survey of Markov decision programming techniques applied to the animal replacement problem. Eur. Rev. Agric. Econ. 21:73–93.

Nielsen, S. S., C. Enevoldsen, and Y. T. Gröhn. 2002a. The *Mycobacterium avium* subsp. *paratuberculosis* ELISA response by parity and stage of lactation. Prev. Vet. Med. 54:1–10.

Nielsen, S. S., Y. T. Gröhn, and C. Enevoldsen. 2002b. Variation of the milk antibody response to paratuberculosis in naturally infected dairy cows. J. Dairy Sci. 85:2795–2802.

Nielsen, S. S., C. Grønbæk, J. F. Agger, and H. Houe. 2002c. Maximum likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53:191–204.

Nielsen, S. S., H. Houe, S. M. Thamsborg, and V. Bitsch. 2001. Comparison of two ELISAs for serological diagnosis of paratuberculosis (Johne's disease) in cattle using different subspecies strains of *Mycobacterium avium*. J. Vet. Diagn. Inv. 13:164–166.

Nielsen, S. S., B. Kolmos and A. B. Christoffersen. 2004. Comparison of contamination and growth of *Mycobacterium avium* subsp. *paratuberculosis* on two different media. J Appl. Microbiol. 96:149-153.

Nielsen, S. S., and N. Toft. 2002. Optimisation of the validity of ELISA and faecal culture tests for paratuberculosis: Selection of population or correction by population characteristics? Pages 400-405 in Proceedings from the Seventh International Colloquium on Paratuberculosis, June 11-14, 2002, Bilbao, Spain. pp. 400–405.

Nielsen, S. S., N. Toft, B. M. Bibby, and A. Nielsen. 2003. Estimation of seroprevalence of paratuberculosis with mixture models including covariates. Pages 3 in Proceedings of the 10th International Symposium on Veterinary Epidemiology and Economics, Vina del Mar, Chile.

Nilsson, D., and A. R. Kristensen. 2002. Markov LIMID processes for representing and solving renewal problems. Pages 97-104 in Proceedings of the First European Workshop on Sequential Decisions under Uncertainty in Agriculture and Natural Resources. Toulouse, France, September 19-20.

Paper IV

Age-specific characteristics of ELISA and fecal culture for purpose specific testing for paratuberculosis

Nielsen SS, Toft N

Published in
Journal of Dairy Science 2006, 89, 569–579

Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis

S. S. Nielsen, N. Toft

Department of Large Animal Sciences, The Royal Veterinary and Agricultural University, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark

ABSTRACT

Paratuberculosis is a chronic infection, and animals are not equally affected by it. Therefore, diagnostic tests that are able to detect different stages of the infection are needed for objective decision making. A longitudinal study was carried out to describe the ability of 2 tests to predict 2 conditions in dairy cattle: "infection" and "infectious," exemplifying 2 different purposes of testing. "Infection" is the term of choice for certification and eradication purposes, and "infectious" is more relevant for control purposes. In the study period of 3 yr, repeated sampling of milk (n = 23,219) and feces (n = 8,832) was performed. A total of 1,985 Danish dairy cows provided material for the study. Milk samples were analyzed for antibodies using an ELISA, and fecal samples were analyzed for mycobacteria by culture. A reference test to correctly classify cattle antemortem does not exist; thus, "infection" and "infectious" were defined by repeated testing using one test as the condition to be detected by the other test. Fecal culture responses were evaluated against antibody status, and ELISA responses were evaluated against detected bacterial shedding. The results of this study indicate that the ability of both tests to detect "infection" increases almost linearly from 2 to 5 yr of age, whereas the ability of both tests to detect "infectious" is not affected by age. Purpose-specific tests are required to appropriately interpret and use test results for management of paratuberculosis, and relevant covariates, such as age, should be included when possible.

INTRODUCTION

Paratuberculosis in cattle is a chronic infection caused by *Mycobacterium avium* ssp. *paratuberculosis* (Map; Chiodini et al., 1984). Infection is usually assumed to take place in calves, and susceptibility to infection decreases as the animal ages (Taylor, 1953; Larsen et al., 1975). Efficient control of infection involves breaking the transmission routes through changes in management practices. Procedures that protect susceptible calves from infectious cows are described as keys in controlling paratuberculosis in infected herds, whereas test-and-cull strategies alone apparently do not decrease the prevalence (Groenendaal et al., 2002). Even so, diagnostic tests are used and may be a significant aid in risk management in infected herds to detect animals that contribute most to the bacterial load in the environment and in milk fed to calves. Diagnostic tests may also be used for certification and other purposes.

The diagnostic accuracy of a test is usually described by its diagnostic sensitivity (the probability that a test is positive given that infection is present) and the diagnostic specificity (the probability that a test is negative given that infection is not present). Underlying the sensitivity and specificity measure is an implicit assumption of an infection that is either present or absent in each animal at a level of interest for particular purposes and resulting decisions. The "infection" condition must reflect the purpose of testing as defined by the decision maker. This requires an understanding of the progressive stages of infection and

accuracies for various testing schemes so that appropriate actions can be initiated subsequent to a positive test result.

In general, purpose-specific objectives could be 1) certification, 2) confirmation of clinical disease, 3) detection of infected animals, 4) detection of animals that are about to become an economic burden, and 5) detection of infectious animals that are shedding large amounts of bacteria to the environment, thus representing a risk to susceptible animals. Decisions subsequent to a test-positive result could be culling, treating the test-positive animals as high-risk animals, or confirmatory testing to provide further information.

For this study, it is assumed that paratuberculosis develops in 3 stages: 1) infection, which generally affects calves; 2) infectious (i.e., shedding infective doses of Map in feces or milk), which follows infection after some variable period of time; and 3) end-stage disease with production losses. Obviously, the strategy of the decision maker influences the optimal choice of which of these 3 stages constitutes a truly positive or infected animal. To establish freedom from disease, information at stage 1 is necessary, but to control the infection to minimize losses, stage 2 might be adequate. Hence, it is necessary to define the purpose of the testing to establish the underlying condition that the test is intended to identify; only then may the diagnostic properties of the test be evaluated.

Two commonly used diagnostic techniques are 1) culture of Map from fecal samples or fecal culture (FC), and 2) detection of antibodies using ELISA. These tests are imperfect for detection of infection, primarily because of a long, and probably variable, incubation period. However, it is uncommon to assess the diagnostic information both relative to the purpose of testing and the chronicity of infection. Furthermore, it is known that other covariates, such as age, influence the test response, especially for the ELISA. Thus, as age of the animal is often readily available, such information should be used when evaluating a diagnostic test.

Although all infected animals must be assumed to be infectious to some degree, some animals may be more infectious than others. During the cell-mediated immune responses, some control of infection is still maintained (Stabel, 2000), and infectiousness is expectedly kept relatively low. During the subsequent humoral immune responses, the infectiousness is expected to be higher than during the cell-mediated immune responses. Thus, it can be assumed that the appearance of antibodies is a predictor of infectiousness. Therefore, an antibody (AB) test, such as an indirect ELISA, would make a good choice given that it can predict highly infectious animals.

The objective of this study was to describe 2 tests, a FC test and an indirect milk ELISA, for their ability to predict 2 conditions: "infection" and "infectious," when the tests are used as screening tests adjusted for age as a covariate. The primary focus was describing the probability of detecting the conditions (sensitivity) and secondarily to describe the probability of absence of the conditions (specificity).

MATERIALS AND METHODS

Herds, Animals, and Observations

The sample population consisted of all cows that have had at least one calf and were present in 8 Danish dairy herds at any time during the study period from January 2000 to March 2003. In all 8 herds, Map had been isolated using FC of samples. During the study period, milk samples were obtained 11 times/yr from all lactating cows in each herd through the Danish milk recording system. Cows that were not lactating did not contribute milk samples

on a given sampling date. Four times per year, fecal samples were collected from all lactating and nonlactating cows in each herd.

Table 1. Description of 8 herds included in the study: Herd size, housing system, production level and apparent prevalence of *Mycobacterium avium* subsp. *paratuberculosis*.

Herd	Housing system	No. of cow yr[1]	Kg. FCM[2] /cow yr[1]	Prevalence (%) of fecal culture positive cows at 1 sampling			Age distribution at 1st sampling in herd			Age distribution at 1st calving		
				Min.	Median	Max.	Min.	Median	Max.	Min.	Median	Max.
1	Tie stall	66.0	8246	1.6	9.9	17.2	2.0	3.9	9.2	1.8	2.2	3.0
2	Bed stall	105.6	10060	0	2.2	5.3	1.8	3.6	10.8	1.8	2.2	3.4
3	Tie stall	119.9	7943	0	2.1	5.9	2.2	3.6	7.3	1.7	2.4	3.6
4	Tie stall	71.2	7383	0	5.0	9.1	2.1	3.8	11.0	1.9	2.3	3.0
5	Bed stall	82.7	7313	0	10.4	18.1	2.1	3.4	7.9	1.7	2.1	3.5
6	Bed stall	260.2	8314	4.4	7.5	13.7	2.2	4.0	12.9	1.7	2.4	3.5
7	Tie stall	68.7	5922	0	12.5	34.8	1.9	3.7	11.0	1.8	2.1	3.1
8	Bed stall	81.5	8138	0	4.8	10.8	2.1	3.9	8.1	1.9	2.4	3.7

[1]In the period Oct. 1, 1999 to Sept. 30, 2000: 1 cow yr is equivalent to 365 cow d.
[2] FCM = fat corrected milk

A summary of information regarding milk production and herd structure is given in Table 1. Information on milk production, breed, and age was obtained from the Danish Cattle Database. Date of birth was missing from 10 cows that were excluded from the study. The herds consisted of 6 different breeds, including crossbred cows. Breeds represented by few animals including 11 Red Danish, 1 Finnish Ayrshire, and 1 Old Danish were excluded from the study. The distribution of the breeds of the remaining cows was 1,430 Danish Holsteins, 435 Danish Jerseys, and 120 crossbreds. These 1,985 cows contributed a total of 23,219 milk samples and 8,832 fecal samples.

The number of samples per cow varied because of the observational nature of the study; i.e., all cows present at the date of first sampling constituted the starting sample. New cows entered at first calving, and older cows left when they were sold, culled, or died. Culling could have occurred because of FC results, as those results were communicated to the farmer. However, culling based on a positive FC was not a common practice among the farms. About 15% of the fecal samples had contaminated culture tubes, and an accurate culture result could not be obtained. Cows from which only contaminated tubes were obtained were excluded from the study and were not included in the 1,985 cows.

The distribution of milk samples per cow was as follows: minimum = 1 sample per cow, median = 16 samples per cow, and maximum = 31 samples per cow. The distribution of fecal samples per cow was as follows: minimum = 1 sample per cow, median = 4 samples per cow, and maximum = 13 samples per cow. The distribution of milk samples per year of age is summarized in Table 2, and the distribution of fecal samples per year of age is summarized in Table 3.

Table 2. Distribution of 23,219 milk samples collected from 1985 Danish dairy cows. Observations are cross-tabulated by age the milk samples were obtained with the age the cow tested positive for *Mycobacterium avium* subsp. *paratuberculosis* in fecal culture (FC). Values in parentheses are mean OD_C-values in the group.

Age when cows were first time detected as FC^+	Yr of age at milk sampling					Total
	<2	2 to 3	3 to 4	4 to 5	>5	
Never	64	4976[1]	5484[1]	3396[1]	4712	18,632
	(-0.01)	(-0.08)	(0.02)	(0.09)	(0.08)	
2 to 3 yr of age	16	555[1]	300[1]	54	0	925
	(-0.01)	(0.32)	(0.67)	(0.90)	(-)	
3 to 4 yr of age	1	469[1]	724[1]	334[1]	85	1613
	(-0.11)	(0.00)	(0.31)	(0.49)	(0.44)	
4 to 5 yr of age	1	126	338[1]	390[1]	173	1028
	(-0.08)	(-0.14)	(0.00)	(0.44)	(0.62)	
5 to 6 yr of age	0	11	79	143[1]	328	561
	(-)	(-0.15)	(-0.07)	(0.13)	(0.37)	
>6 yr of age	0	0	7	50	403	460
	(-)	(-)	(-0.10)	(-0.06)	(0.29)	
Total	82	6137	6932	4367	5701	23,219

[1]) Observations used in the analyses, corresponding to samples obtained the 2 yr following the year a cow was detected FC^+. Samples obtained from cows < 2 yr of age were also excluded

Table 3. Distribution of 8832 fecal samples collected from 1985 Danish dairy cows. Observations are cross tabulated by the age at which fecal samples were obtained with the age of the cow when testing positive for antibodies (AB) to *Mycobacterium avium* subsp. *paratuberculosis* in ELISA

Age when cow first time became AB^+	Yr of age at fecal sampling					Total
	<2	2-3	3-4	4-5	>5	
Never	27	1273[1]	1292[1]	744[1]	1060	4396
<2 yr of age	4	0	0	0	0	4
2 to 3 yr of age	3	292[1]	138[1]	25	2	460
3 to 4 yr of age	1	483[1]	613[1]	202[1]	26	1325
4 to 5 yr of age	1	212	437[1]	363[1]	128	1141
5 to 6 yr of age	0	52	152	217[1]	352	773
>6 yr of age	0	1	25	87	620	733
Total	36	2313	2657	1638	2188	8832

[1] Observations used in the analyses, corresponding to samples obtained the 2 yr following the year a cow was detected AB^+. Samples obtained from cows < 2 yr of age were also excluded

Diagnostic Procedures

Milk samples were tested for presence of antibodies to Map using a milk AB ELISA. All samples were tested in duplicate, and samples differing in corrected optical density (OD_C) by > 0.1 were retested. The ELISA test and its performance have previously been described (Nielsen, 2002; Nielsen and Toft, 2002; Nielsen et al., 2002b, c), with an *M. avium* antigen and absorption with *Mycobacterium phlei*, both obtained from Allied Monitor (Ames, Fayette, MO). Nonspecific reactions are likely to occur, and the diagnostic sensitivity is affected by the age of the animal. The sensitivity and specificity of the test will be affected by the chosen detection limit. As ELISA response, the OD_C was obtained by subtracting the optical density value of a negative control from the mean optical density value of a given milk sample. These OD_C values were used as the outcome in the statistical analyses.

Fecal samples were processed as follows: 1 to 2 g of feces was decontaminated with 10 mL of NaOH solution followed by centrifugation at 1,300 x g. The supernatant was discarded, and the remaining material was dissolved in 5% oxalic acid with 0.1% malachite green. After a further incubation step, the material was centrifuged, and the supernatant was discarded. Neomycin sulfate and amphotericin B were added to the solution and incubated. After incubation and mixing, 3 to 4 drops of solution were applied on each of 4 tubes of medium. The fecal samples were cultured on Löwenstein-Jensen medium before July 2002 and on Herrold's egg yolk medium from August 2002 to the end of the testing period. Samples collected in July and August 2002 were tested in both media, and Herrold's egg yolk medium was more sensitive. Hence, it was decided to change and use the more sensitive medium. All positive cultures were confirmed for presence of the IS*900* sequence in PCR. Further descriptions of the methods and the comparison of the 2 media are given in Nielsen et al. (2004).

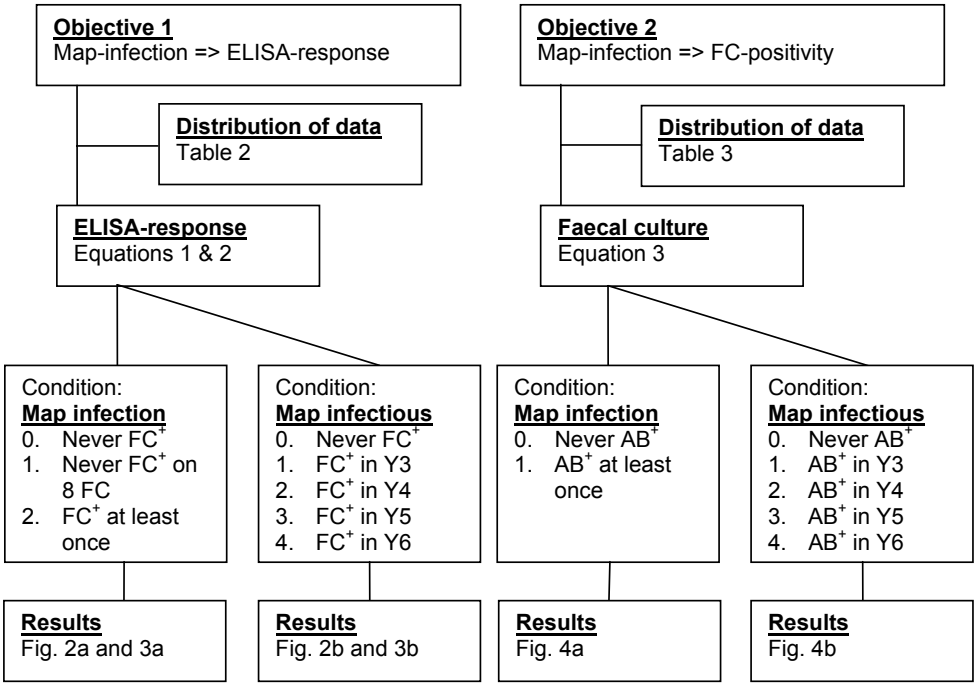


Figure 1. Schematic representation of statistical analyses. FC = Fecal culture; AB^+ = Antibody-positive based on milk ELISA; Y = Yth year of life

Data Preparation

Data control was done by comparing the cow identity of a sampled cow with the information in the Danish Cattle Database to determine whether the cow had actually been in the herd at testing.

Figure 1 is a schematic representation of the statistical analyses. "Objective 1" refers to the investigations of ELISA as the diagnostic test, and "Objective 2" refers to FC as the diagnostic tests. The FC and ELISA were evaluated against the conditions: "Map infection"

and "Map infectious." Because there is no antemortem reference test that can correctly determine the true status of an animal, the following definitions were used:

For Objective 1 (evaluation of the ELISA) the condition “infection” was considered present in any cow that tested FC^+ at least once in her life. She could be classified as “infectious” as follows: Never FC^+; FC^+ in third yr of life (2 to 3 yr of age); FC^+ in fourth yr of life; FC^+ in fifth yr of life; FC^+ in sixth yr of life; or FC^+ in seventh yr of life or later.

For Objective 2 (evaluation of the FC) the condition “infection” was considered present in any cow whose moving average of 2 consecutive OD_C-values was > 0.3. These animals were also referred to as antibody (AB) positive. The moving average at a given age was calculated by taking the average of the present OD_C-value and the OD_C-value on the previous test of the cow. She could be classified as “infectious” as follows: Never AB^+; AB^+ in third yr of life; AB^+ in fourth yr of life; AB^+ in fifth yr of life; or AB^+ in sixth yr of life; or AB^+ in seventh yr of life or later.

Statistical Analyses

Cross-tabulations showing the distributions of test-positive samples (FC^+ or AB^+) obtained by age group and by infectious group (age group at which an animal became test-positive) contributed to the basic descriptive statistics. For descriptive purposes, mean ODC values were calculated for the various groups.

In the following analyses, cows contributed samples in the age group in which they were found FC^+/AB^+ and 2 age groups later, as shown in Table 2 (for analyses of the ELISA response) and Table 3 (for the analyses of the FC).

Predictions of the AB response as a function of age were estimated by nonparametric regression of ODC values as a function of age, using cubic-smoothing splines (Hastie and Tibshirani, 1991) using the GAM procedure in SAS (Version 8.2, SAS Inst., Inc., Cary, NC). The predictions were performed for all observations in 2 groups (FC^+ and FC^-) in which the underlying condition was "infection" and were divided into groups depending on which age they turned FC+ or when the underlying condition was "infectious." The model used for each of these groups was as follows:

$$OD_C = \beta_0 + S(Age) \text{ , per FC-group} \quad (1)$$

where OD_C was the corrected OD-value from the ELISA;

β_0 was the baseline value of the OD_C; and

S(Age) was the smoothing function of age in yr for each of the FC-groups.

To assess the specificity of the ELISA an additional analysis was performed. In this analysis, the non-infected cows were defined as cows with ≥ 8 negative FC obtained over a 2-yr period. Samples from the 2-yr period were excluded. Only samples from about the first year were included.

Subsequent to analysis of the ODC-response on a continuous scale, the OD_C was dichotomised at a cut-off of 0.3, which is the recommended laboratory cut-off (where OD_C > 0.3 are positive). Data were analyzed with the following model:

$$g(P(OD_C > 0.3)) = \log\frac{P(OD_C > 0.3)}{1 - POD_C > 0.3)} = \beta_0 + S(Age) \text{ , per FC-group} \quad (2)$$

where $P(OD_C > 0.3)$ was the probability of testing positive in the ELISA at cut-off 0.3;

β_0 was the baseline probability of testing positive in ELISA; and

S(Age) was the smoothing function of age in years in each FC-group.

As with the OD_C on a continuous scale, a separate analysis was done with cows which had at least 8 negative FC over a 2-yr period. Samples from the 2-yr period were excluded. Only samples from about the first year were included.

A similar model was used for predictions of the probability of testing positive at any time in FC. The model was:

$$g(P(FC^+)) = \log\frac{P(FC^+)}{1-P(FC^+)} = \beta_0 + S(Age)\text{, per } AB^+\text{-group} \qquad (3)$$

where $P(FC^+)$ is the probability of testing positive in FC;

β_0 was the baseline probability of testing positive in FC; and

$S(Age)$ was the smoothing function of age in yr in each AB-group.

The predictions for FC^+ were done for all observations divided into 2 groups: AB^+ and AB^- with the underlying condition being "infection", and also when divided in groups depending on the age at which they turned "infectious".

The effect of change in FC method was assessed by an extension of Equation 3, where culture method was included as a parametric term, which could be either the old or the new method.

RESULTS

Table 2 gives the distribution of observations used for the analyses of OD_C values, cross-tabulated by actual age group at sampling and the age group in which the animal first became FC^+, as well as the mean OD_C value in each of the groups. Table 3 gives the distribution of observations used for analyses of FC response; the cross-tabulation now includes the age at which an animal first turned AB+. A total of 302 cows contributed 2,485 samples to the specificity assessment of the ELISA. These cows had a minimum of 8 negative FC over a 2-yr period, and only samples obtained before that period were included.

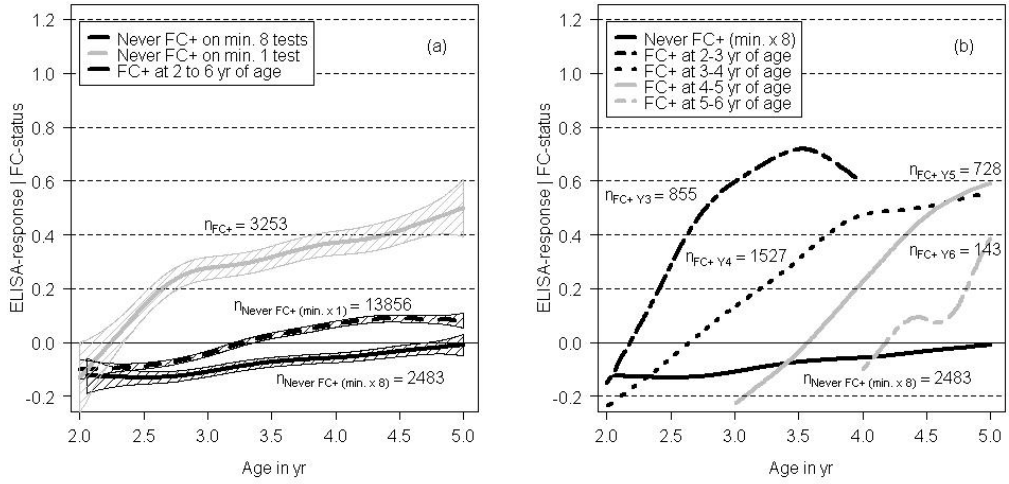


Figure 2. Predicted corrected optical density (OD_C) value in *Mycobacterium avium* subsp. *paratuberculosis* antibody ELISA used for milk samples, estimated for groups of cows divided on the basis of the age at which they tested positive in fecal culture (FC) if ever. The number of observations (n) which contributed to each prediction is given in the diagram. In Figure 2a, the

predictions are shown for 3 groups of cows: Never FC^+ (minimum 1 FC), never FC^+ (minimum 8 FC), and FC^+ sometime from 2 to 7 yr of age. In Figure 2b, the predictions are shown for 5 groups, in which FC^+ cows were divided into groups depending on when they became FC^+ (Never; 3^{rd}; 4^{th}; 5^{th} or 6^{th} yr of life). The shaded areas around each graph in Figure 2a are the 95% confidence bands; these are omitted in Figure 2b to improve readability.

In Figure 2a, the predicted OD_C values are shown for 2 groups of cows, one group never becoming FC^+ and the other group becoming FC+ at some time during 2 to 6 yr of age. These are the predictions of the ELISA response for "infection." In Figure 2b, the latter group was further divided into 5 groups, one for each age group in which cows became FC^+. These are predictions of the ELISA-response for being "infectious." The estimated probabilities of testing ELISA-positive are shown for the same groups in Figures 3a and b, subsequent to the dichotomization of the ELISA response.

Cows tested before 2 yr of age did not show indications of having antibodies, as the graphs for infected and noninfected cows have the same starting point at 2 yr of age (Figure 2a). The average OD_C value increases steeply from 2 to 2.5 yr of age. After 2.5 yr of age, the increase is less steep. A high OD_C value will typically be in the range 1.0 to 1.4, although OD_C values > 2.0 were observed for 120 samples in the data. Thus, the average covers a wide variation of OD_C values. This variation can be examined further by analyzing the data in the age groups, based on when cows became FC positive (Figure 2b). The average OD_C in each of these groups increases more steeply than the average in Figure 2a, thus emphasizing the relation between immune response and bacterial shedding. The almost parallel appearance of the graphs (Figure 2b) is supporting evidence that the age of OD_C positivity is better explained by the age of FC positivity than the age itself; i.e., infectiousness is predicted by the appearance of an immune response.

In Figures 3 and 4, the sensitivity and specificity estimates obtained in the present study can be read on the vertical axis, although they are referred to there as probabilities given a condition. The sensitivity of the test for a given age is read directly as the probability for the positive groups, whereas the specificity given age is read as 1 – the probability for the group in which the condition was never found. This condition may be interpreted as infection for Figures 3a and 4a and infectious for Figures 3b and 4b.

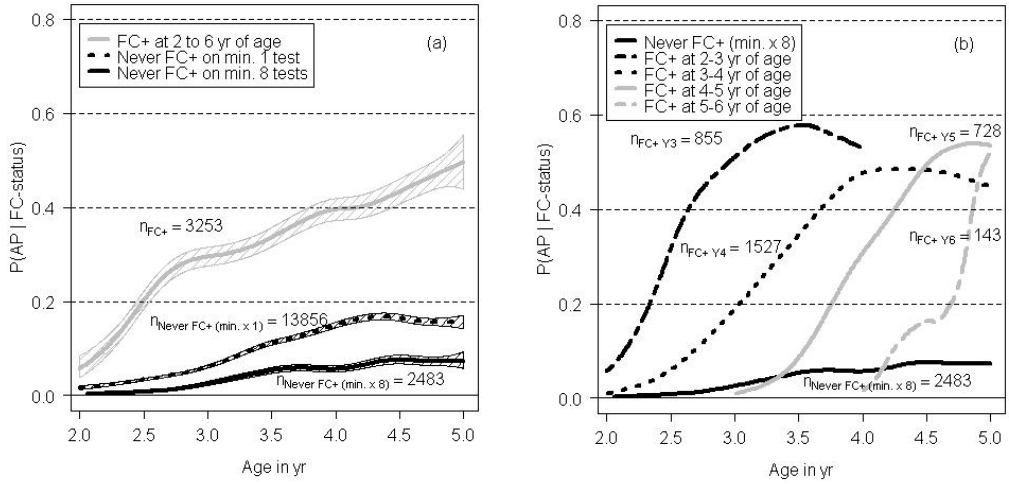


Figure 3. Estimated probability of testing positive in an ELISA for detection of antibodies to *Mycobacterium avium* subsp. *paratuberculosis*, for cows grouped according to the year in which they first tested positive in fecal culture (FC^+). The number of observations (n) contributing to each predicted curve is given in the diagram. In Fig. 3a, the predictions are shown for 3 groups of cows: Never FC^+ (minimum 1 FC), never FC^+ (minimum 8 FC), and FC^+ sometime during the age of 2 to 7 yr. In Fig. 3b, the predictions are shown for 5 groups, in which FC^+ cows were divided into groups depending on when they became FC^+ (never; 3rd; 4th; 5th or 6th yr of life). The shaded areas around each graph in Fig. 3a are the 95% confidence bands; these are omitted in Fig. 3b to improve readability.

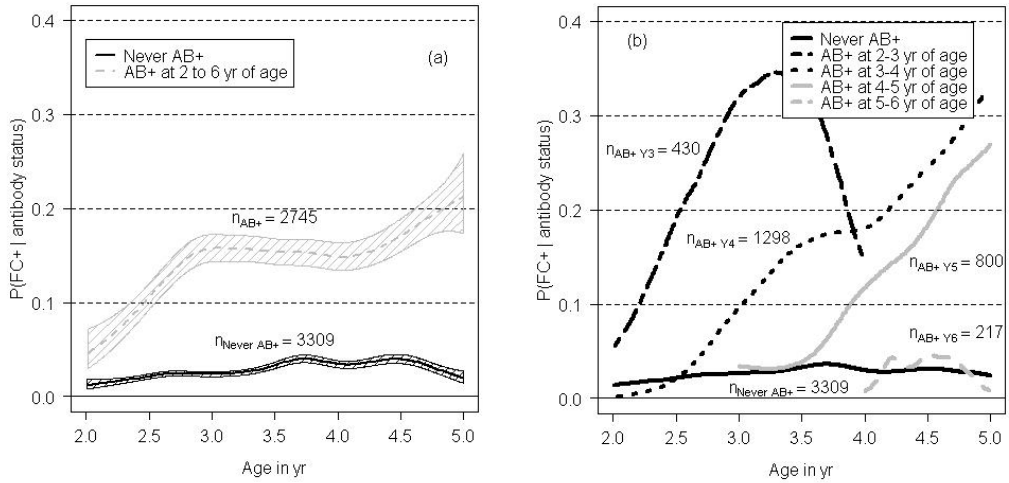


Figure 4. Estimated probability of testing positive in a fecal culture test for detection of *Mycobacterium avium* subsp. *paratuberculosis*, for cows grouped according to the age at which they first tested antibody positive (AB^+) in an antibody ELISA. The number of observations (n) contributing to each predicted curve is given in the diagram. In Fig. 4a, the predictions are shown for 2 groups of cows: Never AB^+, and AB^+ sometime between 3rd to 6th yr of age. In Fig. 4b, the predictions are shown for 5 groups, in which AB^+ cows were divided into groups depending on when they became AB^+ (3rd; 4th; 5th; or 6th yr of age). The shaded areas around each graph in Fig. 4a are the 95% confidence bands.

The sensitivity of ELISA for prediction of "infection" increases from 0.06 at 2 yr of age to 0.50 at 5 yr of age. The specificity based on samples obtained from cows before 8 negative FC decreases from 0.997 to 0.93 in the same age span, where the specificity is calculated as 1 – the probability defined in the graph (Figure 3a). The sensitivity of the ELISA to predict current MAP shedding, or "infectiousness," in a specific FC^+ age group (cows FC+ between 2 to 3 yr of age) is increasing from 0.06 to 0.58 at 3.6 yr of age. Similar increases in sensitivities are observed in the other groups of cows, stratified into the age group when they became FC^+ (Figure 3b). Had the age groups been more narrow; e.g., split into half-year intervals, the sensitivity had increased to 0.7 for cows FC+ in age range 2 to 2.5 yr of age and 0.9 for cows FC^+ in age range 2.5 to 3.0 yr of age (data not shown). However, because the data underlying these estimates were sparse, the uncertainty associated with the estimates requires that caution is used when interpreting the results. In Figure 4a, the estimated probability of testing FC^+ ($P(FC^+|$AB status)) is shown for 2 groups of cows: those that never tested AB^+ and those that tested AB+ some time during years 3 to 6 of life, corresponding to the "noninfected" and the "infected" group, respectively. In Figure 4b $P(FC^+|$AB status) is shown for 5 groups based on the age group in which they became AB^+, if ever. These are predictions of the FC for being "infectious." The sensitivity of FC increases with age. At 2 yr of age, the sensitivity is 0.05, and at 5 yr of age, the sensitivity is 0.21. The "specificity" was in the range 0.964 to 0.986 with higher specificity for young cows.

The effect of culture method was insignificant averaged over all groups of AB^+ ($P = 0.22$). Hence, for ease of interpretation, the results given were averages of both methods.

DISCUSSION

Generally, there is a linkage between ELISA and FC. In this study, this linkage appears to be independent of age. Whereas there is a linkage between the 2 tests independent of age, the sensitivity of both tests still increases with age, and the specificity of ELISA apparently decreases with age. This is the first study describing the test responses of ELISA and FC as a function of age in a large population of naturally Map-infected cattle. The underlying condition for evaluating test responses was varied to allow purpose-specific test interpretation. The following assumptions were used: 1) Map infection takes place in calves, 2) Map infection lasts for life, and 3) some infected cows are not as infectious as others; i.e., the incubation period is variable and cow-specific. Purpose-specific test evaluation and interpretation with inclusion of the important covariate age could be an important extension in the use of imperfect tests in the control of paratuberculosis. The present study can be seen as a first step toward a purpose-specific test evaluation. There are a number of issues, which should be considered in the interpretation of the results, and a number of approaches can be used to improve the interpretation.

Disease Definition, Sensitivity and Specificity

Given a unique disease definition (the underlying condition), the ability of a test to detect this condition can easily be characterized. However, for paratuberculosis, a suitable, unique disease definition does not exist. For decision makers seeking to eradicate Map from a herd or region, the appropriate definition of paratuberculosis varies in that decision makers may merely be wishing to reduce economic losses or they may be hopeful of using the test results for appropriate management of infectious cows.

A frequent approach for the evaluation of tests for detection of Map infection is the selection of a sample population, for which infection status is determined postmortem. Such a population is often older than the population that is to be evaluated by the test subsequently. It may also be selected based on criteria, which favor agent-detecting methods (such as FC), because bacteria detected are deemed more indicative of infection than the finding of an immune response. Selection biases are difficult to avoid, because the test scheme chosen as the reference or gold standard essentially defines what constitutes absence and presence of infection (Nielsen and Toft, 2002).

Latent-class methods, which do not presume the definitive status of an animal, can be used, as exemplified in Nielsen and Toft (2002). Currently, these methods do not take into account the purpose of testing, which still need to be addressed. However, latent class methods may provide a promising alternative for the analyses carried out in the present study.

In this study, the repeated testing of animals by FC was used as a reference for the ELISA test, and the repeated testing of milk samples using ELISA was used as a reference for the FC. These are not perfect reference methods, but will underestimate the true sensitivity of each method because the low sensitivity of any reference method will falsely classify some test-positive as false-positive even though they are actually true positives that are not detected by the reference method if the reference method has a low sensitivity. However, to simplify the discussion, we will assume that when used for classification into "infected" or "infectious" the relevant reference method is perfect. Thus, for evaluation of ELISA, repeated FC is a perfect reference method and vice versa. The conditions "infected" and "infectious" were assumed to be the same irrespective of reference method.

Sensitivity and Specificity Given Disease Definition: Infection

Figures 2a, 3a, and 4a illustrate the test responses as a function of age when the underlying condition is infection. Given the aim is to detect an infected animal, these figures show which test responses can be expected with increasing age. For ELISA, the sensitivity increases almost linearly from 0.06 at 2 yr of age to 0.50 at 5 yr of age (Figure 3a), whereas the specificity decreases from 0.997 to 0.93 as the age of the tested animal increases. This apparent drop in specificity is likely to be an artifact caused by the low sensitivity of FC (Figure 4a), although some of this potential artifact can be removed by testing more frequently with FC. The average sensitivity of FC increases from 0.05 at 2 yr of age to 0.21 at 5 yr of age. The specificity is high (0.964 to 0.984). A true false-positive is a positive test response from an animal that is not infected. It is possible, although unlikely, that a bacterium detected is not Map, as the cultured bacteria were confirmed with IS900 PCR. However, false-positives may be caused by pass-through, in which cows consume Map without becoming infected and lead to a possible false-positive test reaction. Another possibility is that ELISA has not detected cows that were shedding bacteria subsequent to Map infection. The conclusion at this stage must be that the ELISA is more sensitive for detection of Map infection than FC, but at an early age, it is still low. There may be a risk of obtaining many false-positives, but it is unlikely that they are all attributable to false-positive reactions of the ELISA. Many could be due to the low sensitivity of FC.

Sensitivity and Specificity Given Disease Definition: Infectious

The results shown in Figures 2b, 3b, and 4b are the test responses when the underlying condition is infectious. First, the ELISA response, on average, for a cow that is detected FC^+ in her third year of life will also test positive within that year (Figures 2b and 3b), although initially, the sensitivity is only 0.06 at 2 yr of age, but at 3 yr of age, it has increased to 0.58 (Figure 3b). The pattern is similar for cows found positive in FC in their fourth and fifth year of life: initially, the sensitivity is low, but steeply increases within the following year. The increases in sensitivity are almost parallel for the different age groups. This indicates that a positive OD_C is coherent with a positive FC and vice versa. Thus, the cow may transfer to the infectious stage any time, regardless of age of the cow (although still in the interval 2 to 5 yr of age). For FC, the sensitivity is also increasing from 0.06 at 2 yr of age to 0.35 at 3 yr of age for cows that are AB^+ in their third year of life. These results indicate that there is a good relationship between being infectious and AB^+ and that both the FC and the ELISA may detect some of these animals. However, the ELISA has a greater sensitivity than the FC, although to some degree it may be at the expense of specificity, particularly for older animals. The apparent loss in specificity is likely caused by the generally low sensitivity of the FC method. This can partly be deduced from Figure 2b. A cow testing FC^+ in her fifth year of life does not have AB before 3 to 3.5 yr of age, whereas the average of the FC^- cows is higher. These would be expected to be similar if the FC^- cows were all correctly classified. Given that some FC^- animals are actually infected, the sensitivity of ELISA is underestimated. The magnitude of this underestimation could easily be 0.05 to 0.10, corresponding to the apparent drop in specificity.

If the OD_C is used as an approximation of AB production, it is observed that AB are generally present in low levels at 2 to 3 yr of age, whereas higher levels are not reached before the animals are 4 to 5 yr of age. This is consistent with previous findings (Nielsen et al., 2002a). Detected fecal shedding is maximum around 4 to 5 yr of age, with a maximum $P(FC^+|AB^+) = 0.17$. The maximum at 4 to 5 yr of age is consistent with the findings by Kalis et al. (1999). If AB was used as the gold standard in a test evaluation, the $P(FC^+|AB^+)$ would have corresponded to a diagnostic test sensitivity of 0.17. This is consistent with what has been obtained using methods in the absence of a gold standard (Nielsen and Toft, 2002; Nielsen et al., 2002c), in which sensitivity estimates in the range 0.2 to 0.4 were obtained. The lower sensitivity found here reflects the misclassification of false-positives introduced because of the lack of specificity of the ELISA.

Cows that are never AB+ have a probability of 0.02 to 0.04 of testing FC^+. Hence, most FC^+ cows will, at some point in time, test positive in ELISA. Those that do not become AB^+ are potentially passive carriers of Map without ever being infected, as described by Sweeney et al. (1992). Another explanation is that they are unable to produce AB, perhaps because they were infected in uteri while the immune system was developing. Hence, Map could produce persistent infections similar to those occurring with bovine virus diarrhea infections. A last possible explanation is censoring, i.e., FC^+ cows have not been kept long enough after first shedding has been detected, and therefore, production of AB simply has not begun. The finding suggests that, in most cases, the ELISA will at some time detect infection, but not necessarily before the cow becomes infectious to other cattle.

Concluding Remarks

Discerning between "infection" and "infectious" can be highly relevant for a decision maker, as the infected cow may prohibit a declaration of freedom from infection, but this cow may never become an economic burden. Hence, if a farmer has no intention of declaring freedom from infection, he may only want to focus on detecting infectious animals to reduce spread of infection. Infectious animals can be detected with ELISA with a fair sensitivity in all age groups even with a single testing. Repeated tests can be required to further characterize the infection state, e.g., whether or not the cow will experience production losses. Infected animals cannot be detected at younger ages, but this is less relevant if the objective only is to detect infectious animals. Thus, a clear aim of the testing strategy should be defined before testing. Given this strategy, an advisor may provide more specific estimates of the probability of detecting the condition, whether this specific condition is infection or infectiousness.

The nature of this study was descriptive; thus, to simplify the presentation of the results, some elements have been given less attention. Consequently, some of the results should be interpreted carefully. For some groups, the number of observations is low, and generally, the right ends of the curves are more uncertain that the rest because of the censoring of old cows. To simplify the results and because of the computational complexity, uncertainty estimates have generally been excluded from the presentation. However, 95% confidence limits are given in Figure 2a, 3a, and 4a. It should be noted that correlation between samples from the same cows has not been addressed, and single observations are not independent. Finally, the predictions presented are average values, which do not take into account the variation in the response to infection of individual cows as previously described (Nielsen et al., 2002b). This variation may be reflected in variation among herds. Variation attributable to herd could not be assessed because of the lack of statistical power. However, there were indications that cows of Herds 2 and 7 did not, on average, yield similar responses as cows in the other herds.

To handle some of the issues raised here, other techniques, such as time-to-event analyses, would be more appropriate, with the events being AB^+ and FC^+. Also assessing the relations between the 2 events would be important. However, the purpose of the current studies was primarily descriptive and should be considered as such.

Diagnostic sensitivity and diagnostic specificity are test properties normally characterizing diagnostic tests in control and eradication programs. With a clear disease definition and consensus on actions following diagnosis, sensitivity and specificity are adequate and informative characterizations of a test. However, some decision makers require tests for conditions other than those optimal in eradication programs, for example. The chronic nature of paratuberculosis and the lack of perfect diagnostic tests is a continuous challenge in the interpretation of the test results obtained in any testing. Chronicity affects the sensitivity in that progression of infection increases the sensitivity. In a given population, progression may not only be determined by a fixed incubation time. It may be influenced by management factors (e.g., feeding strategy) that lead to variation in incubation times. Yet, the fixed incubation time approach can provide us with some information, e.g., by inclusion of age as factor in interpretation of the test results. Another factor is the definition of actions that are made subsequent to a diagnosis, when potential production losses may require one action whereas transmission of Map may call for another, with different requirements at different prevalences.

CONCLUSIONS

The results of this study indicate that the ability of both ELISA and FC to detect "infection" increases almost linearly from 2 to 5 yr of age. The ability of the 2 tests to detect "infectious" is not affected by age. Purpose-specific test evaluations should be conducted to appropriately interpret and use test results for management of paratuberculosis, and relevant covariates should be included when possible. For detection of infection, age is a relevant covariate.

ACKNOWLEDGMENTS

The Danish Cattle Federation financed this study, partly via Kongeåprojektet, partly via CEPROS-II/7.

REFERENCES

Chiodini, R. J., H. J. van Kruiningen, and R. S. Merkal. 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74: 218-262.

Groenendaal, H., M. Nielen, A. W. Jalvingh, S. H. Horst, D. T. Galligan, and J. W. Hesselink, 2002. A simulation of Johne's disease control. Prev. Vet. Med: 54, 225-245.

Hastie, T. J., and R. J. Tibshrinani. 1991. Generalized additive models. Chapman & Hall, London, England.

Kalis, C. H .J., J. W. Hesselink, E. W. Russchen, H. W. Barkema, M T. Collins, and I. J. R. Visser. 1999. Factors influencing the isolation of *Mycobacterium avium* subsp. *paratuberculosis* from bovine faecal samples. J. Vet. Diagn. Invest. 11: 345-351.

Larsen, A. B., R. S. Merkal, and R. C. Cutlip. 1975. Age of cattle related to resistance to infection with *Mycobacterium paratuberculosis*. Am. J. Vet. Res. 36: 255-257.

Nielsen, S.S. 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subsp. *paratuberculosis* in dairy cattle. J. Vet. Med. B 49: 384-87.

Nielsen, S.S., and N. Toft. 2002. Optimisation of the validity of ELISA and faecal culture tests for paratuberculosis: Selection of population or correction by population characteristics? *In:* Juste, R. A., M. V. Geijo, and J. M. Garrido (Eds). Proceedings from the Seventh International Colloquium on Paratuberculosis, June 12-14, 2002, Bilbao, Spain, 400-405.

Nielsen, S.S., C. Enevoldsen, and Y.T. Gröhn. 2002a. The *Mycobacterium avium* subsp. *paratuberculosis* ELISA response by parity and stage of lactation. Prev. Vet. Med. 54: 1-10.

Nielsen, S.S., Y.T. Gröhn, and C. Enevoldsen. 2002b. Variation of the milk antibody response to paratuberculosis in naturally infected dairy cows. J. Dairy Sci. 85: 2795-802.

Nielsen, S.S., C. Grønbæk, J.F. Agger, H. Houe, 2002c. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53: 191-204.

Nielsen, S. S., B. Kolmos, and A. B. Christoffersen, 2004. Comparison of contamination and growth of *Mycobacterium avium* subsp. *paratuberculosis* on two different media. J. Applied Microbiol. 96: 149-153.

Stabel, J. R., 2000. Transistions in immune responses to *Mycobacterium paratuberculosis*. Vet. Microbiol. 77: 465-477.

Taylor, A. W. 1953. Experimental Johne's disease in cattle. J. Comp. Pathol. 63: 355-367.

Sweeney, R. W., R. H. Whitlock, A. N. Hamir, A. E., Rosenberger, and S. E. Herr. 1992. Isolation of *Mycobacterium paratuberculosis* after oral inoculation of uninfected cattle. Am. J Vet. Res. 53: 1312-1314.

Paper V

Transitions in diagnostic tests used for detection of *Mycobacterium avium* subsp. *paratuberculosis* infections in cattle

Nielsen SS

Published in
Veterinary Microbiology 2008, 132, 274–282

Transitions in diagnostic tests used for detection of *Mycobacterium avium* subsp. *paratuberculosis* infections in cattle

S. S. Nielsen

Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark

ABSTRACT

Diagnosis of infections with *Mycobacterium avium* subsp. *paratuberculosis* (MAP) is difficult due to a long incubation period and lack of tests which can accurately predict the future status of animals. Early detection of infectious animals is necessary to reduce transmission of MAP. The objective of this study was to determine the time from first detection of MAP-antibodies in milk ELISA to start of MAP shedding, for animals with various shedding patterns.

An observational longitudinal study was carried out over 3 years. A total of 24,076 milk and 10,074 faecal samples were obtained from 1906 cows and tested using ELISA and FC, respectively. Cows were classified into 5 shedding groups based on their repeated FC: non-shedders (NS; *n* = 1512 cows, 79.3% of total), transient (TS; *n* = 36, 1.9%), intermittent (IS; *n* = 137, 7.2%), low (LS; *n* = 143, 7.5%), and high shedders (HS; *n* = 78, 4.1%).

Results showed that 5% of TS, 30% of IS, 60% of LS and 70% of HS were ELISA-positive at the date of first positive FC, and many HS (28%) and LS (14%) were positive ≥1 year prior to first detection of shedding. FC confirmed shedding within the first year after the positive ELISA in 10% of 328 cows with fluctuating ELISA compared with 35% of 445 cows with the last 2 or more ELISAs positive.

To conclude, MAP-antibodies were generally detected prior to start of bacterial shedding, with difference between the various patterns of shedding, and a positive ELISA was useful for predicting that an animal would subsequently become infectious.

INTRODUCTION

Paratuberculosis in cattle is a chronic infection caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP) (Chiodini et al., 1984). Control programmes have been established in several countries (Kennedy and Nielsen, 2007), because MAP infections can be costly to the cattle industry (Ott et al., 1999). Most programmes include diagnostic testing, although purposes of testing can differ with different objectives of the programmes. The requirements for diagnostic tests ability to differentiate between conditions such as “infectious”, “infected” and “non-infected” would therefore vary from one type of programme to another. Infectious animals make up an immediate risk for spread of MAP, whereas the infected animal constitutes a future risk of spread of MAP. Detection of non-infected animals is of particular importance in certification schemes. Therefore, a decision maker would need to know the accuracy of a test relative to the use of a diagnostic test result, and estimates of sensitivity and specificity should be estimated relative to a given target condition.

Timely identification of infectious animals is the primary purpose of testing in the Danish programme, in order to reduce MAP transmission from these animals to susceptible herd mates (Nielsen, 2007). Sensitivities of diagnostic tests for ante mortem diagnosis of MAP infectious animals have been reported in the range 0.21–0.94, with a median of 0.46

(reviewed in Nielsen and Toft, 2008). Data from the literature do not suggest that one test has a better accuracy than others, but ELISA is an in-expensive alternative to faecal culture (FC). As an example, sampling and test costs for one milk ELISA and one FC were estimated to 3 EUR and 27 EUR, respectively, in a herd with whole-herd testing of 120 cows in 2006 in Denmark (Sergeant et al., 2008).

Infections with MAP are characterised by long incubation periods, which may be of variable length. The prevailing hypothesis on the pathogenesis is that some time after infection, cell-mediated immune responses occur, followed by humoral immune responses (Stabel, 2000). Cows shedding MAP on average have higher concentrations of IgG antibodies (Koets et al., 2001), but the patterns of shedding related to immune responses are poorly described. Because an antibody ELISA test is an indirect measure of infected or infectious status of an individual, the relationship between the detection of antibodies and commencement of shedding is a key factor in designing an effective control program. Shedding of MAP is usually defined based on results from an agent detecting test such as FC in different media or direct PCR. However, various categories of shedding can occur, depending on the number of bacteria shed and pattern of shedding (transient, intermittent or continuous). Transient shedding may be either the result of uninfected cows' ingestion of MAP from the environment with subsequent shedding of MAP in the following days (Sweeney et al., 1992), or it may simply be a typical pattern seen in the transition from the infected to the infectious stage of disease. Intermittent and repeated shedding would be expected to occur as infection progresses. This variable pattern of detection may be due to either intermittent excretion of MAP or excretion of low levels of MAP around the limits of detection of the FC system utilised. There is no specific explanation why ELISA-positive results can first be detected in some animals prior to evidence of shedding and after confirmation in others. Some ELISA-positive animals remain FC negative indefinitely even though they are infected (e.g. confirmed by tissue culture post-mortem).

The primary objective of the current study was to estimate the time relationship between occurrence of antibodies, as determined by ELISA, to first detection of MAP by FC, for cows with various shedding patterns.

MATERIALS AND METHODS

Herds, Animals, and Observations

Data were collected from a non-random sample of 8 Danish dairy herds in the period 11 August 1999 to 12 December 2002. A brief description of the herds is given in Table 1. Faecal sampling was performed four times a year on all cows present in each herd at the time of sampling. Milk samples were collected from all lactating cows via the Danish milk-recording scheme on 11 test-dates per year. The study design was observational; hence each cow provided variable number of samples.

A total of 1906 cows provided both milk (n = 24,076) and faecal samples (n = 10,074). The distributions of number of samples per cow were as follows: (a) milk samples—minimum: 1; 1st quartile: 6; median: 11; 3rd quartile: 18; and maximum 35; (b) faecal samples—minimum: 1; 1st quartile: 2; median: 5; 3rd quartile: 8; and maximum 18. Information on date of birth and breed was obtained from the Danish Cattle Database. The breed distribution of the cows was 1350 Danish Holsteins, 427 Danish Jerseys, 12 Red Danish, 1 Finnish Ayrshire, 1 Old Danish, and 115 crossbreds. The age distribution at 1st testing of the individuals was minimum: 1.74 years; 1st quartile: 2.32 years; median: 2.56

years; 3rd quartile: 3.44 years; and maximum 12.49 years. Age distribution at first sampling was minimum: 2.05 years; 1st quartile: 3.01 years; median: 3.91 years; 3rd quartile: 4.91 years; and maximum 10.70 years.

Table 1. Description of 8 study herds: herd size, housing system, breed distribution and milk production level

Herd	Housing system	Cow years[2]	Kg ECM[1]/ cow year[2]	Cow years[3]	Kg ECM/ cow year[3]	Breed distribution[4]
1	Tie stall	66.0	8152	64.1	8500	97% DH & 3% Cross
2	Bed stall	105.6	10002	112.5	10390	100% DH
3	Tie stall	119.9	7833	155.6	7282	71%DH % & 29% Cross
4	Tie stall	71.2	7368	69.6	7413	52%DH & 45%DJ
5	Bed stall	82.7	7284	122.5	7855	100%DJ
6	Bed stall	260.2	8270	258.9	8222	95%DH & 5%Cross
7	Tie stall	68.7	5910	70.5	5601	97%DJ & 3%Cross
8	Bed stall	81.5	8050	91.8	8211	100%DH

[1]Kg energy corrected milk yield (kg ECM) was estimated from the milk yield control scheme carried out in the herds 11 times per year; [2]For the period Oct. 1, 1999 to Sept. 30, 2000; [3]For the period Oct. 1, 2001 to Sept. 30, 2002; [4]Breed distribution at the beginning of the study. DH=Danish Holstein; DJ=Danish Jersey; Cross=Crossbred

Diagnostic tests

Milk samples were tested for presence of antibodies using the ELISA (Nielsen, 2002) used in the Danish control programme on paratuberculosis (Nielsen, 2007). Sensitivity and specificity estimates for detection of infected animals vary with age and have been described in detail in Nielsen and Toft (2006). All samples were tested in duplicate, and the mean result was used. Samples were retested if the difference in optical density (OD) values was >0.1. The mean OD-value of a sample was corrected for intra-laboratory variation by subtracting the mean OD-value of a negative control. Samples with a corrected OD-value (ODC) of ≥0.3 were deemed positive, as done in the Danish control programme.

Faecal samples collected prior to August 2002 were cultured on Löwenstein–Jensen Medium and samples from September 2002 and onwards were cultured on Herrold's Egg Yolk Medium, while samples from August and September 2002 were cultured on both media. The methods are described in Nielsen et al. (2004). IS900 PCR was carried out on positive cultures to confirm that presumed isolates were MAP. Positive cultures were recorded on an ordinal scale as follows—1+: <10 colonies; 2+: 10–49 colonies; and 3+: ≥50 colonies.

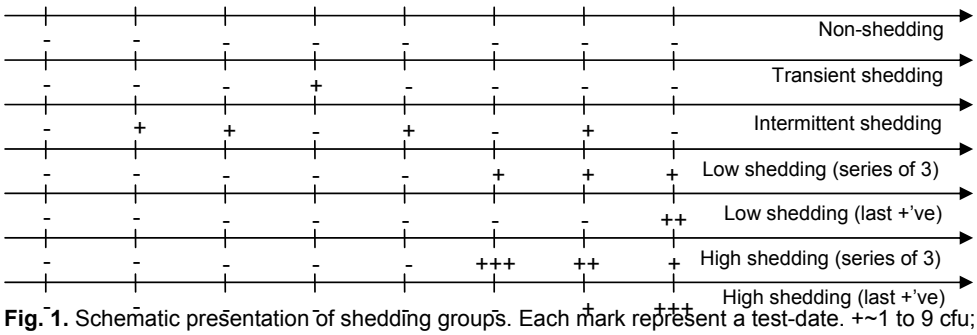


Fig. 1. Schematic presentation of shedding groups. Each mark represent a test-date. +~1 to 9 cfu; ++~10 to 49 cfu; +++~≥50 cfu.

Statistical analysis

Definitions

Cows were grouped into MAP shedding groups based on their FC results as follows (illustrated in Fig. 1):

- Non-shedding. No positive cultures among all bacteriological results from a cow.
- Transient shedding. Last four samples were MAP negative but at least one positive result was recorded previously.
- Low shedding. Cows where all samples were FC-positive after the first positive sample or where the last three samples were positive and none of the samples had ≥50 CFU per test-tube.
- High shedding. Cows fulfilling the same criteria as for low shedding, but at least one culture had ≥50 CFU per test-tube.
- Intermittent shedding. Cows with at least one positive FC, but not meeting the criteria of the other shedding groups.

The date of FC positivity (Time0) was defined for all shedding cows as the date of the first detection of a positive sample. For cows never shedding MAP, Time0 was defined as the median test-date for the individual cow. The distribution of samples within each shedding group is shown in Table 2.

Table 2. Distribution of number of faecal samples per cow, and proportions of faecal culture positive samples per cow in total and after first positive sample, stratified by shedding group

Shedding group[§]	Total no. of samples					% positive samples among all samples					% positive samples after 1st positive sample				
	Min	Q1	Q2	Q3	Max	0%	1-25%	26-50%	51-75%	>75%	0%	1-25%	26-50%	51-75%	>75%
Never	1	5	10	14	18	1512	0	0	0	0	1512	0	0	0	0
Transient	5	8	10	12	16	0	36	0	0	0	0	36	0	0	0
Intermittent	2	5	9	12	16	0	61	60	15	1	0	9	79	45	4
Low	1	5	9	13	16	0	54	42	14	33	0	0	0	2	141
High	1	4	7	10	14	0	5	20	16	37	0	0	1	3	74

Abbreviations: Min=Minimum; Q1, Q2 and Q3=1st, 2nd and 3rd quartile, respectively; Max=maximum

Cows were also divided into antibody categories based on their antibody profiles determined by ELISA. The antibody categories were defined as those in the Danish control programme as follows: (A0) a minimum of two milk samples were available, and all samples had ODC < 0.3 (negative); (A1) only one sample, which was negative; (A2) last three samples were negative, but previously a minimum of one sample had been positive (ODC ≥ 0.3); (A3) last sample was negative, but some previous samples had been positive (fluctuating response); (A4) last sample was positive, but previous samples were all negative; and (A5) last two or more samples were positive (Nielsen, 2007). The date of positivity (Time0) was defined as the date of first positive ELISA.

Descriptive statistics

Proportions of positive faecal samples per animal in total and after first positive sample were calculated within each MAP shedding group (Table 2). For each of these groups, the distribution of ODC-values was also described (Table 3). Proportions of positive faecal

samples in each of the antibody categories overall and after first positive ELISA were also calculated (Table 4).

Table 3. Distribution of ELISA-results (corrected optical density (OD_C) for each shedding group

MAP Shedding group	No. of cows	No. of samples	OD_C				
			Min.	Q1	Q2	Q3	Max.
Never	1,512	18,838	0.00	0.00	0.00	0.06	2.58
Transient	36	785	0.00	0.00	0.00	0.13	2.25
Intermittent	137	2,092	0.00	0.00	0.03	0.41	2.58
Low	143	1,597	0.00	0.00	0.09	0.55	2.81
High	78	764	0.00	0.00	0.20	0.81	2.68
Total	1,906	24,076	0.00	0.00	0.00	0.10	2.81

Abbreviations: Min=Minimum; Q1, Q2 and Q3=1st, 2nd and 3rd quartile, respectively; Max=maximum

Table 4. Proportion of faecal samples and positive faecal cultures (FC) overall and after first positive ELISA within each antibody category

Antibody category	No. of cows	No. of FC	Overall positive FC (%)	Positive FC after 1st positive ELISA (%)
A0: Always ELISA$^-$	932	4134	114 (2.8%)	NA
A1: ELISA$^-$, but only 1 sample tested	53	71	4 (5.6%)	NA
A2: Last 3 samples ELISA$^-$, but previously min. 1 sample ELISA$^+$	100	517	8 (1.5%)	4 (1.9%)
A3: Fluctuating ELISA, last sample ELISA$^-$	263	1664	92 (5.5%)	55 (9.7%)
A4: Last sample ELISA$^+$, previous samples ELISA$^-$	136	674	35 (5.2%)	7 (21.2%)
A5: Last 2 or more samples ELISA$^+$	422	2873	459 (16.0%)	360 (27.7%)

NA: Not applicable

Analytical statistics

The probability of testing positive in ELISA relative to Time = 0 for each shedding group was estimated using a generalised additive model (Hastie and Tibshrinani, 1990) using the nonparametric logistic regression model:

$$\log\frac{P(\mathrm{OD_C}>0.3)}{1-P(\mathrm{OD_C}>0.3)}=\beta_0+S(\mathrm{Time})$$

where $P(OD_C>0.3)$ was the probability of being ELISA-positive, β_0 the baseline probability of testing ELISA-positive, and S(Time) was the smoothing function of the effect of time relative to first positive FC (Time0) within a shedding group.

A similar logistic regression model was used with FC as the dependent variable and S(Time) as the independent variable, for each of the antibody categories. The effect of change of faecal culture method was also assessed. For both models, 95% point-wise confidence bands were estimated. The generalised additive models were estimated using the GAM procedure in SAS v. 9.1 (SAS Institute, Cary, North Carolina, USA).

RESULTS

Distributions of positive FC among all samples and among samples obtained after the first positive FC are summarised in Table 2 for each shedding group. Proportions of FC-positive samples within each antibody category are shown in Table 4, both overall and after 1st

positive ELISA result. The distribution of ODC-values within each shedding group is shown in Table 3.

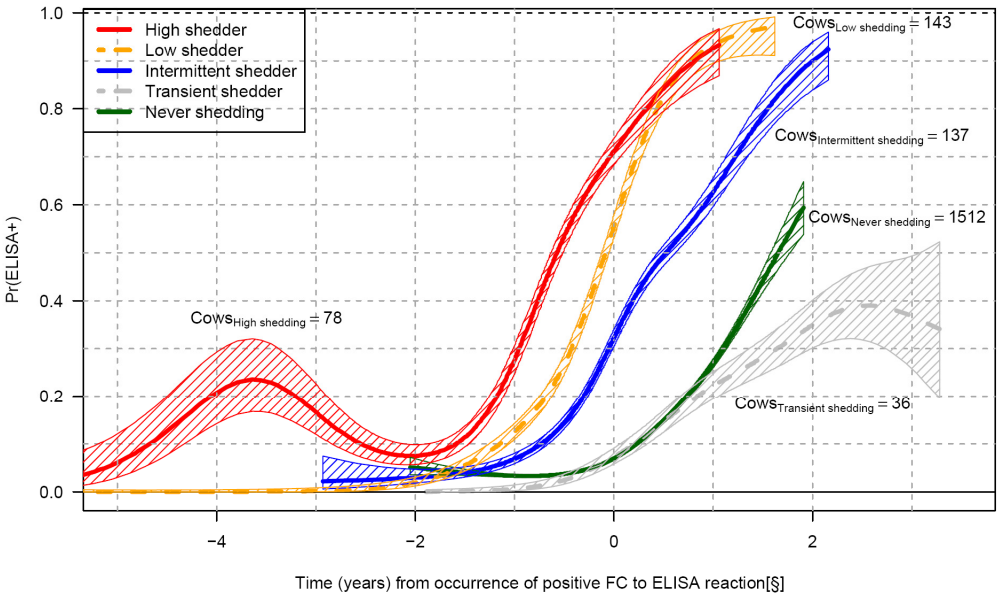


Fig. 2. Probability of testing positive by ELISA relative to time, where "Time" was the date of entering a shedding group. [§]For cows never shedding MAP, Time=0 was defined as the median ELISA test-date. 95% confidence bands are shown for each graph. Graphs only extent to the time for which data was available.

Probabilities of testing positive by ELISA relative to Time0 for each of the shedding groups are shown in Fig. 2. Among "high shedders" an estimated 70% were ELISA-positive on the first detection date of MAP shedding, whereas 60% of cows in the "low shedding" group were ELISA-positive at Time0. For "intermittent shedders", 30% were ELISA-positive at the first date of detected shedding. Among "transient shedders" an estimated 7% were ELISA-positive at Time0 with a significant proportion (40%) of animals ELISA-positive 2–3 years after Time0. The "non-shedder" group showed an estimated 7% ELISA-positive at Time0 and an estimated 27% ELISA-positive at 1 year and 60% at 2 years after the median test-date. Approximately 20% of the cows in the "high shedder" group were positive 3–4 years prior to detection of shedding, but this estimate was associated with some uncertainty compared to other estimates. The probability of being ELISA-positive among cows in the "low shedder" group was 0% in 3–5 years prior to start of shedding, and cows in all shedding groups other than the "high shedders" had low probabilities (<4%) of being ELISA-positive 2–3 years before shedding was first detected.

There was no difference in probability of testing FC-positive with the two different culture methods ($p = 0.38$). Probabilities of testing positive in FC for cows in various antibody categories are shown in Fig. 3, with Time0 the testing date on which the animal first tested positive. A cow repeatedly positive by ELISA (antibody category A5) would test positive by FC in 20% of cases at Time0. The optimal time for confirming the ELISA test result by FC would be 9 months after the milk sample had been obtained when an estimated 35% of A5

cows would be positive. Cows with only the last ELISA-positive (A4 category) showed 17% FC-positive at Time0 rising to 21%. Cows with fluctuating ELISA-responses (A3 category, e.g. negative–positive–negative–positive–negative results) had an estimated 8% probability of testing FC-positive at Time0 and maximum 12% probability of testing positive approximately 6 months later. The 95% confidence bands for estimates (Fig. 3) for all antibody categories were generally quite narrow except for those in A4 category (last ELISA-positive) where there was a very marked spread of the confidence bands after Time0 due to a limited number of observations after Time0 in this category.

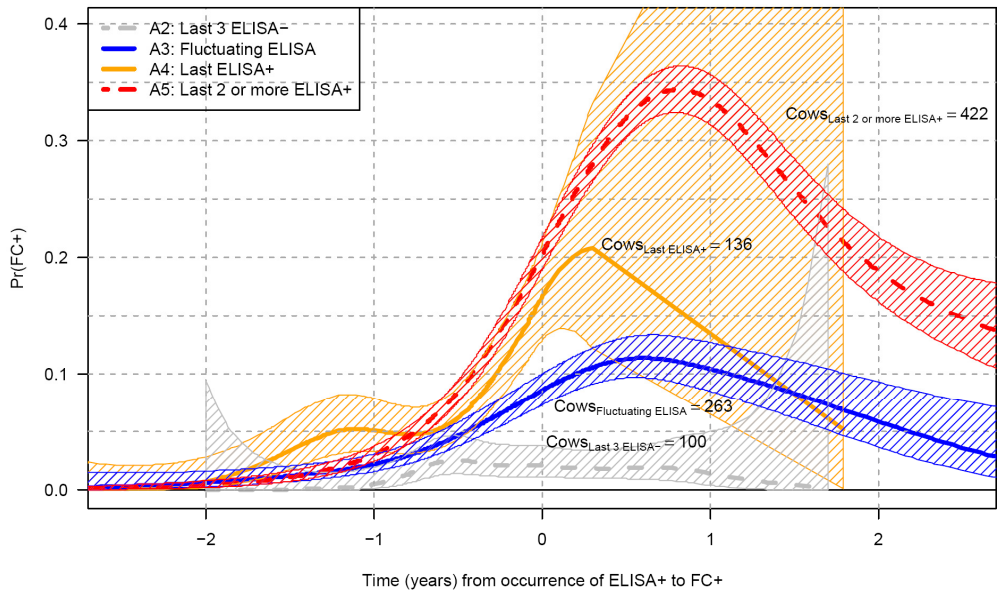


Fig. 3. Probability of testing positive by FC for cows with different ELISA profiles. "Time" = 0 was the date of 1st positive ELISA for a given cow. 95% point-wise confidence bands are shown with shading around the point estimates.

DISCUSSION

The results in this report show that the ELISA has a high level of predictive ability that an animal is or will become infectious. Antibodies are generally detectable prior to MAP shedding, but a positive ELISA result is not necessarily an indicator of shedding on the test-date. The almost parallel lines in Fig. 2 show an interesting progression of MAP infection. Cows with "transient shedding" generally do not have a positive ELISA reaction prior to Time0, while larger proportions of cows are ELISA-positive with progression from "transient" over "intermittent" and "low" to "high" shedding. Thereby, a continuous progression of MAP infection appears to occur, and ELISA results are linked to deterioration of shedding. Due to the complex nature of the interaction between the immune response and MAP, and the long incubation period, division into "infectious", "infected" and "non-infected" is not straightforward, neither based on the immune response, nor based on the detection of MAP using an agent detecting test such as FC. The advantage of FC is that detection of MAP is a direct measure of MAP shedding on the test-date. However, the amount of MAP shed need not constitute an infective dose, or the shedding may cease after a short shedding period

and re-occur later, as demonstrated by Lepper et al. (1989). The advantage of detecting an immune response is that ELISA is less expensive than FC and results can be available faster. Frequent testing using ELISA is therefore more attractive to farmers. Antibody profiles based on frequent testing may subsequently suggest whether there is a high or low risk of shedding until next test-date. The diagnosis provided is a probability diagnosis, because final proof of shedding is not provided. Cows with repeated positive ELISA are more likely to be shedding MAP in the near future than are fluctuating immune responses (Fig. 3). Detection of animals that are or will become infectious in the time-span until the next test-date is pivotal in a control programme to avoid transmission. Subsequent decisions on management of test-positive animals depend on a variety of factors such as milk quota, prevalence of test-positives, availability of replacement animals, stage in lactation, clinical signs (of MAP infection or other conditions of priority in the herd), milk yield, facilities for management of the potentially infectious animals in facilities where they do not make a risk for susceptible animals, etc. The farmer's ultimate goal of intervention is also important, while different goals often require different decisions. Possible decisions could include (a) follow-up testing; (b) culling based on one ELISA result only; (c) “isolation”, i.e. measures should be taken to avoid transmission of MAP from test-positive to susceptible animals; and (d) do nothing. The ideal evaluation of such decision processes should be through simulation models (e.g. Kudahl et al., 2007), where cost and impact of false-positive and false-negative test-results can be accounted for. Costs related to misclassification can vary greatly from one herd to another.

Follow-up testing with ELISA, FC or other tests is an option in the decision making process. The test-history may be part of the decision process, and multiple previous positive test results may suggest that re-testing is not needed. Two or more ELISA-positive results suggest that a cow has a higher probability of shedding MAP, but the basis for making a decision will vary from one decision maker to another. It should be noted that probabilities presented in this paper should be read as one diagnostic test result at a given time point, as the observations were not independent. Hence, the cumulative probability leading up to a time point is greater than the single estimated probability at that time point but less than the sum of the probabilities for the series of time points. Most cows shedding MAP will be ELISA-positive at some point in their lifetime (Nielsen and Ersbøll, 2006). Should a decision maker wish to confirm MAP shedding predicted by an ELISA-result, the highest chance of confirmation by FC is 1/2–3/4 year after the ELISA was performed (Fig. 3).

ELISA-results may in some cases be false-positive. However, considering that most cows in the pre-shedding phase in the present study were ELISA-negative, it is unlikely that antibodies to non-MAP infections caused the high probabilities of testing ELISA-positive seen in cows that were never shedding MAP. Data presented here suggest that specificity of the ELISA was >97.5%, based on the early ELISA-responses in all shedding groups, excluding ELISA-responses of high shedders (Fig. 2). Repeated testing with ELISA could increase specificity, although not to 100%. A high specificity is primarily required for herd-classification purposes, but while FC is not appropriate for confirmatory purposes in infected but non-infectious animals, other methods could be used for herd-classification, e.g. the Rogan–Gladen estimator for prevalence estimation (Rogan and Gladen, 1978) as exemplified in Sergeant et al. (2008).

The probability curves (Fig. 2) suggest that occurrence of antibodies was a good indicator of progression of MAP infection, but many non-shedding animals will test-positive. The high proportions of ELISA-positive animals among non-shedding and transient shedding

animals suggest that many of these animals were in an early phase of infection. “Transient shedding” may not be “passive shedding” in the meaning described by Sweeney et al. (1992), where un-infected cows fed high dosages of MAP shed MAP after a few days. It appears that “transient” as defined in the current study may merely be a precursor of intermittent, low and high shedding. The occurrence of the phenomenon “passive shedding” cannot be confirmed with the current study design.

The current study has a number of limitations. Firstly, the herds may not be representative of the cattle population. A complete random inclusion of animals would be preferred, and the animals should then be followed for their entire lifetime with follow-up post-mortem examinations to verify their infection status. Such a design was not possible, partly for practical reasons, partly because it would be extremely expensive. Secondly, the ELISAs available world-wide appear to vary greatly in sensitivity and specificity (Nielsen and Toft, 2008), and few ELISAs have been appropriately compared. Therefore, caution should be made if conclusions from this study are transferred to other ELISAs. Lastly, caution should be made when interpreting point estimates in the graphs, when wide 95% confidence intervals are present. Wide confidence bands result primarily for two reasons: great variation in data or low number of observations. For example, all confidence bands in Fig. 2 and Fig. 3 were wide in the right side of the graphs due to the generally low number of observations at the later time points, an intrinsic aspect in the analysis where the cows were sampling occurred over a 3 year time frame and first positive results could occur at any point within that time frame. However, a wide confidence band for the high shedding group around Time = −4 years (Fig. 2) was seen due to great variation in test-responses in combination with low number of samples. The interpretation of the graphs was best at points with narrow confidence intervals.

The results of this study are the first observational longitudinal study in a large and naturally infected population of dairy cattle demonstrating the relation between test-responses of ELISA and FC, and occurrence of antibodies and bacterial shedding. The results are reasonably consistent with the theory described regarding the pathogenesis (Stabel, 2000), and with an experimental study (Lepper et al., 1989). However, Lepper et al. (1989) also detected shedding of MAP at early ages (<1 year), shedding that waned to subsequently reappear at older ages. The current study did not include young animals and could therefore not confirm this finding from the animals experimentally infected. More studies with longitudinal study designs should be carried out to verify the results presented here.

Implementation of the results from this study in control, surveillance and eradication programmes would require studies on technical and economical effects, to determine which test-strategy should be preferred. Sergeant et al. (2008) demonstrated that whole-herd ELISA would be the most cost-effective test-strategy for surveillance with the prevalences, herd-sizes and test-costs seen in Denmark, but this may be different in other scenarios. In control programmes, it would also be important to determine if farmers would respond to the recommendations given, while testing is of no use if appropriate actions are not made to a test-result.

CONCLUSIONS

The results in this study indicate that antibodies to MAP generally occur prior to shedding of MAP, although transient or intermittent shedding can occur prior to the occurrence of

antibodies. Progression of MAP infection is generally accompanied by occurrence of antibodies. Positive ELISA-results can only be confirmed with FC in a fraction of animals, because time from occurrence of antibodies to shedding of MAP varies greatly.

ACKNOWLEDGEMENTS

The project was supported by the Danish Directorate of Food, Fisheries and Agro Business (DFFE, J.nr. 3412-06-01550) and the Danish Cattle Federation.

REFERENCES

Chiodini, R.J., van Kruiningen, H.J., Merkal, R.S., 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74, 218-262.

Hastie, T. J., Tibshrinani, R.J., 1991. Generalized additive models. Chapman & Hall, London, England.

Kennedy, D.J., Nielsen, S.S., 2007. Report from the first IDF ParaTB Forum. Bulletin of the International Dairy Federation. 410, 3-7.

Koets, A.P., Rutten, V.P., de Boer, M., Bakker, D., Valentin-Weigand, P., van Eden, W., 2001. Differential changes in heat shock protein-, lipoarabinomannan-, and purified protein derivative-specific immunoglobulin G1 and G2 isotype responses during bovine *Mycobacterium avium* subsp. *paratuberculosis* infection. Infect. Immun. 69, 1492-1498.

Kudahl, A.B., Østergaard, S., Sørensen, J.T., Nielsen, S.S., 2007. A stochastic model simulating paratuberculosis in a dairy herd. Prev. Vet. Med. 78, 97-117.

Lepper, A.W.D., Wilks, C.R., Kotiw, M., Whitehead, J.T., Swart, K.S., 1989. Sequential bacteriologic observations in relation to cell-mediated and humoral antibody responses of cattle infected with *Mycobacterium paratuberculosis* and maintained on normal or high iron intake. Aust. Vet. J. 66, 50-55

Nielsen, S.S., 2007. Danish control programme for bovine paratuberculosis. Cattle Practice, 15, 161-168.

Nielsen, S.S., Ersbøll, A.K., 2006. Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows. J. Dairy Sci. 89, 4557-4566.

Nielsen, S.S., Toft, N.. 2006. Age-specific characteristics of ELISA and fecal culture for purpose specific testing for paratuberculosis. J. Dairy Sci. 89, 569-579.

Nielsen, S.S., Toft, N., 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques. Vet. Microbiol. 129, 217-235.

Nielsen, S.S., Kolmos, B., Christoffersen, A.B., 2004. Comparison of contamination and growth of *Mycobacterium avium* subsp. *paratuberculosis* on two different media. J. Appl. Microbiol. 96, 149-153.

Ott, S.L., Wells, S.J., Wagner, B.A., 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40, 179-192.

Rogan, W.J., Gladen, B., 1978. Estimating prevalence from the results of a screening test. Am. J. Epidemiol. 107, 71-76.

Sergeant, E.S.G., Nielsen, S.S., Toft, N., 2008. Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds. Prev. Vet. Med. 85, 92-106.

Stabel, J.R., 2000. Transitions in immune responses to *Mycobacterium paratuberculosis*. Vet. Microbiol. 77, 465-473.

Sweeney, R.W., Whitlock, R.H., Hamir, A.N., Rosenberger, A.E., Herr, S.A., 1992. Isolation of *Mycobacterium paratuberculosis* after oral inoculation of uninfected cattle. Am. J. Vet. Res. 53, 1312-1314.

Paper VI

Time to occurrence of drop in milk production in cows with various paratuberculosis antibody profiles

Nielsen SS, Krogh MA, Enevoldsen C

Published in
Journal of Dairy Science 2009, 92, 149–155

Time to the occurrence of a decline in milk production in cows with various paratuberculosis antibody profiles

S.S. Nielsen, M. A. Krogh, and C. Enevoldsen

Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark

ABSTRACT

Infection with *Mycobacterium avium* ssp. *paratuberculosis* (MAP) in dairy cattle often results in reduced milk production and premature culling. Some test-positive animals can live for years without being affected by infection, whereas others are test negative when they die from the infection. Our objective was to describe the deviation in milk production of cows with various MAP antibody profiles compared with their repeatedly test-negative herdmates in the same parity. Data were obtained from herds participating in the Danish control program on paratuberculosis, for which 4 annual MAP antibody ELISA of individual cows were performed per herd per year. A total of 136,489 ELISA results from 38,998 dairy cows in 64 herds were used along with 484,285 test-day records on energy-corrected milk (ECM) yield. Cows were divided into 6 antibody groups based on their repeated milk ELISA results: A0) repeated ELISA negative; A1) ELISA negative, but only once; A2) ELISA negative on the last 3 tests, but with 1 previous positive result; A3) ELISA negative on the last test, but with 1 or more previous positive results; A4) last sample was ELISA positive, but all previous were negative; and A5) at least the last 2 samples were ELISA positive. The expected test-day kilograms of ECM by herd and parity were estimated for cows in antibody group A0. Deviations from expected milk production were then assessed for cows in the other antibody groups relative to the time of the first test-positive ELISA result (D 0). Cows in groups A2, A3, and A5 produced approximately 0.5 kg of ECM/d more than cows in group A0 at 300 d before D 0. Cows in group A4 had a decline in milk production from d 300 before D 0, with daily milk production reduced by 5 kg of ECM at 200 d after D 0. Milk production of cows in group A5 was reduced by 2.5 kg of ECM at 300 d after D 0 compared with 300 d before D 0, whereas cows in groups A2 and A3 produced 0.5 kg of ECM more than cows in group A0. The conclusions of the study were that 1) increasing the number ELISA tests increases the predictive value of ELISA for inference on milk production losses, 2) a combination of ELISA with assessment of observed milk production may be a valuable tool for decisions on culling, and 3) the declines in milk production attributable to MAP occurred over a long time period, and may not be realized by the herd manager without more advanced management tools such as the model proposed here.

INTRODUCTION

Paratuberculosis is a chronic infection caused by *Mycobacterium avium* ssp. *paratuberculosis* (MAP). Control or certification programs are established in several countries (Kennedy and Nielsen, 2007), emphasizing the interest of the industry in controlling MAP infections. A primary reason is the economic losses that infection can cause in a herd. The major losses are incurred through reduced milk production and premature culling, with the latter resulting in increased cow replacement costs.

Some discrepancies exist in the literature concerning how much milk production is affected. Milk production was 20 and 17% less in the last lactation compared with the previous lactation among culled cows with clinical symptoms and among cows that were positive in a serum antibody ELISA with MAP infection confirmed by histology, respectively, in a Dutch study (Benedictus et al., 1987). Cows with a strongly positive serum ELISA result produced 1,364 kg less 305-d mature-equivalent milk (Lombard et al., 2005), and cows with positive milk ELISA produced 457 kg less milk per lactation (Hendrick et al., 2005) than ELISA-negative cows. Johnson et al. (2001) reported an increase (P = 0.11) in milk production among serum ELISA-positive cows compared with ELISA-negative cows. These studies were all cross-sectional, with diagnostic procedures carried out only at 1 time point per animal. Wilson et al. (1993) divided cows in a herd into MAP-infected and noninfected based on repeated fecal culture (FC). They found that FC-positive, first-parity cows in the beginning of lactation (100 DIM) produced 1.6% more milk compared with FC-negative cows, whereas FC-positive cows late (>100 DIM) in the second, third, and fourth lactation produced 2.9, 8.2, and 8.4% less milk, respectively, compared with their FC-negative herdmates. These results suggest that animals that become test positive initially have a greater milk production, but later the test-positive animals produce less milk (equal to steeper lactation curves). Nevertheless, Wilson et al. (1993) did not assess the time from when an animal tested positive relative to when the decline in milk production occurred, and the study design was not appropriate for making inferences about when a production loss occurred and whether infected or test-positive animals had a greater milk yield early in the infection. To our knowledge, no studies have been performed to assess the period from test positivity to the occurrence of milk production losses, which may explain the discrepancies in results reported.

Because milk production losses associated with MAP infections can be one of the major motivators for controlling the infection, assessment of the time from a positive test to expected milk production losses would be needed for optimal timing of culling a test-positive animal. Results of the above-mentioned studies suggest that stage of infection affected the occurrence and magnitude of milk production losses. Nevertheless, losses need not be linked to the occurrence of antibodies in serum (Johnson et al., 2001; Hendrick et al., 2005) or to bacterial shedding (Wilson et al., 1993). Humoral immune reactions with the occurrence of antibodies generally characterize the progression of MAP infections, but some antibodies can be detected in the early stages of infection (Koets et al., 2001). Therefore, the antibody profile of a cow could be a better predictor of milk production losses than individual antibody measurements. The objective was to describe the time from detection of antibodies to MAP (as detected by a milk ELISA) to the occurrence of declines in milk production for cows with various antibody profiles.

MATERIALS AND METHODS

Data Collection

Data from the Danish control program on paratuberculosis (Nielsen et al., 2007) and the Danish milk recording scheme were extracted from the Danish Cattle Database and used for the analyses. Of the approximately 4,600 dairy herds in Denmark, 1,219 participated in the program on the date (May 8, 2008) of data extraction. Program herds fulfilling the following criteria were included: 1) more than 200 cows were present in the herd on the date of data extraction, 2) herds were enrolled in the Danish milk recording scheme and had 11 annual

milk production recordings, and 3) more than 90% of the cows in each herd were Danish Holsteins. These criteria were set to make the predicted milk yields of cows without MAP within a given herd as robust and precise as possible. Test-day records were excluded from cows with missing information on age at first calving, non-Holstein cows, when test days were >305 DIM, and when test dates were >3 yr old. The resulting data set consisted of 484,285 test-day records from 38,998 cows.

Herds in the MAP control program were tested 4 times yearly by using an in-house milk antibody ELISA (Nielsen, 2002). The milk samples tested were the same as those obtained in the milk recording scheme for assessment of milk yield, fat, protein, and SCC, and were a composite sample from the morning and evening milking. The test was based on a commercially available *Mycobacterium avium* ssp. *avium* purified protein derivative-antigen (Allied Monitor, Fayette, MO), and both IgG1 and IgG2 were detected by the test. Test responses 0.3 optical density corrected (ODC) values were considered positive, but because of repeated testing, cows could be classified into 6 antibody groups based on their ELISA profile, rather than being evaluated based on a single test value only: A0) repeated ELISA negative based on a minimum of 2 samples; A1) ELISA negative, but only 1 sample was available; A2) ELISA negative on the last 3 samples, but with 1 previous positive result; A3) ELISA negative on the last test, but with 1 or more previous positive results, usually with interchangeable positive and negative reactions ("antibody fluctuator"); A4) the last sample was ELISA positive, but all previous were negative; and A5) at least the last 2 samples were ELISA positive. These 6 antibody groups were defined so that all cows were classified rule based. We used the same terminology and grouping as in Nielsen (2008), in which the time relationship between occurrences of MAP antibodies to shedding of MAP was studied. Cows with positive ELISA reactions were defined as belonging to the specific antibody groups on the date of the first positive result (D 0). Cows that shifted from one antibody group to another were classified based on the last result obtained, but the date of the first positive ELISA was retained as D 0. From the 38,998 cows, 136,489 test results were available: 7,694 cows with 1 test result, 7,027 cows with 2 test results, 7,526 cows with 3 test results, 6,712 cows with 4 test results, 4,621 cows with 5 test results, and 5,424 with 6 or more test results. The data included 23,098 cows in antibody group A0, 6,541 cows in antibody group A1, 728 cows in antibody group A2, 2,913 cows in antibody group A3, 2,832 cows in antibody group A4, and 2,886 cows in antibody group A5.

Statistical Analyses

The statistical analyses were carried out in a 3-step process, in which the average kilograms of ECM yield were described for cows in parity 1, 2, and >2, and average OD_C values were described for cows in each of the antibody groups. The true prevalence was estimated based on the test prevalence, age of the individual, and age-specific estimates of sensitivity and specificity. The estimator described by Sergeant and et al. (2008; available at http://parafree.vetinst.dk/parafree/content.php?page=parafree) was used for the estimations. The predicted kilograms of ECM were then estimated for each cow in a given parity in a given herd for all cows that were repeatedly negative in ELISA (cows in antibody group A0). Finally, this expected milk yield was used to estimate deviations in the milk yield of cows of other antibody groups. These analyses were performed as described below.

Descriptive Statistics. Test-day milk yield was recorded 11 times yearly in routine milk recordings, at which time kilograms of milk, percentage of fat, and percentage of protein were determined. The test-day ECM (ECMT) yield was calculated by using the formula

kg ECMT = (kg milk x (383 x fat% + 242 x protein% + 780.8))/3140 (1)

For each of the 3 parity groups, 1, 2, and >2, the mean, median, first quartile, and third quartile kilograms of ECMT and OD_C were calculated. Subsequently, kilograms of ECMT were estimated as a function of DIM for each parity group by using a generalized additive model (Hastie and Tibshrinani, 1990) using a B-spline smoother with 10 degrees of freedom with PROC GAM in SAS, version 9.1.3 (SAS Inst. Inc., Cary, NC). The resulting average lactation curves were used to determine the shape of the lactation curves for the estimations of expected milk yield.

Average antibody profiles for cows in each antibody group were described by using the same approach as for ECMT, with OD_C as a function of time relative to D 0. For cows in antibody group A0, D 0 was the median between the first and last test dates. The median test date was used because that date would be in the middle of the study period, and a uniform fixed point was needed. For cows in antibody group A1, only 1 observation was available per cow, and these cows were excluded because all observations would be at D 0.

Prediction of Kilograms of ECMT in Antibody Group A0. The lactation curves suggested a peak milk yield between 40 and 65 DIM, and lactation curves were fitted similar to those described by Bennedsgaard et al. (2003), where peak milk yield was 60 DIM. The predicted milk yield for cows in antibody group A0 was estimated by using the 3-level model for separate parity strata, 1, 2, and >2:

$$ECM_{ijk} = \beta_0 + \beta_1 DIMun60_{ijk} + \beta_2 DIM60_{ijk} + \beta_3 AC1_{jk} + \beta_4 Twin_{jk} + \beta_5 Calving_{jk} + \beta_6 Season(Year)_{jk} + \varepsilon_{ijk} \quad (2)$$

where $\beta_0 = \beta_{000} + v_{00k} + \mu_{0jk}$, $\beta_1 = \beta_{1000} + v_{100k} + \mu_{10jk}$, $\beta_2 = \beta_{2000} + v_{20k0} + \mu_{20jk}$ and $v_{00k} \sim N(0, \tau_{00k})$, $\mu_{0jk} \sim N(0, \tau_{0jk})$, $v_{100k} \sim N(0, \tau_{100k})$, $\mu_{10jk} \sim N(0, \tau_{10jk})$, $v_{20k0} \sim N(0, \tau_{20k0})$, $\mu_{20jk} \sim N(0, \tau_{20jk})$ and $\varepsilon_{ijk} \sim N(0, \sigma_{ijk})$.

ECM_{ijk} was the kg ECM on the i^{th} test-day of the j^{th} cow in the k^{th} herd; $DIMun60_{ijk}$ was the i^{th} DIM of the j^{th} cow in the k^{th} herd for DIM 1 to 60. For DIM > 60, DIM60 takes the value 0; $DIM60_{ijk}$ was the i^{th} (DIM-60)/245 of the j^{th} cow in the k^{th} herd for 60 DIM to 305 DIM. For values < 60 DIM, DIM60 takes the value 0; $AC1_{jk}$ was the age at 1st calving for the j^{th} cow in the k^{th} herd (this effect was only included for 1st parity cows); $Twin_{jk}$ was effect of the j^{th} cow giving birth to a twin in the k^{th} herd; $Calving_{jk}$ was effect of calving problems of the j^{th} cow in the k^{th} herd based on the recordings of the calving made by the farmer. Calving was classified as yes or no, where yes was difficult calvings and veterinary assisted calvings; $Season(Year)_{jk}$ was nested effect of the season (1: January to March; 2: April to June; 3: July to September; 4: October to December) in the a study year (2005; 2006; 2007 or 2008) on the j^{th} cow in the k^{th} herd.

The variable β_0 can be separated into an overall mean (β_{000}) that represents the average across-herd milk yield at 60 DIM and a contribution from the individual herds (v_{00k}), and a contribution from the j^{th} cow within the k^{th} herd (μ_{0jk}). The variable v_{0k} was the mean ECM in the kth herd; μ_{0jk} was the mean ECM of the j^{th} cow in the k^{th} herd; β0 was the overall mean of ECM at the intercept; β1, β2, and β3 were the fixed linear regression coefficients of DIM60, DIMun60, and AC1, respectively; β4, β5, and β6 were class effects of the Twin, Season(Year), and Calving, respectively; v_{2k} was the random linear regression

coefficient of $DIM60_{ijk}$; μ_{2jk} was the random linear regression coefficient of $DIM60_{ijk}$; and μ_{2jk} ε_{ijk} was the random residual component assumed independent, identically distributed normal, $N(0, \sigma^2)$.

Deviation from Predicted Kilograms of ECM for Cows from Various Antibody Groups. The predicted kilograms of ECM of cows in antibody group A0 from the same parity in the k^{th} herd was used as the expected kilograms of ECM of a given cow in the same parity and in the same herd. Also included in the predictions were the remaining covariates described in equation [2]. The deviation in kilograms of ECM from the expected milk yield was calculated for cows in antibody groups A2 to A5. The D 0 was used to calculate the time from the ECM test date to a positive ELISA test: T_{Diff} = milk recording test date – D 0.

Deviations of the kilograms of ECM were then estimated as a function of T_{Diff}. The model used for each of these groups was

$$\text{Deviation kg ECM} = \beta_0 + \beta_1 T_{Diff} + \beta_2 T_{Diff}^2 + \beta_3 T_{Diff}^3 + \beta_4 T_{Diff}^4 + \beta_5 T_{Diff}^5 + \varepsilon \quad (3)$$

The results were visualized by plotting the deviation in kilograms of ECM relative to T_{Diff} for each antibody group.

RESULTS

The parity-specific distributions of test-day kilograms of ECM and OD_C-values from ELISA are given in Table 1. Averages of parity-specific kilograms of ECM as a function of DIM are shown in Figure 1. Average ELISA profiles for antibody groups A0 and A2 to A5 are shown in Figure 2. The estimated distribution of within-herd true prevalences was as follows: minimum: 0%; first quartile: 1%; median: 9%; third quartile: 20%; and maximum: 52%.

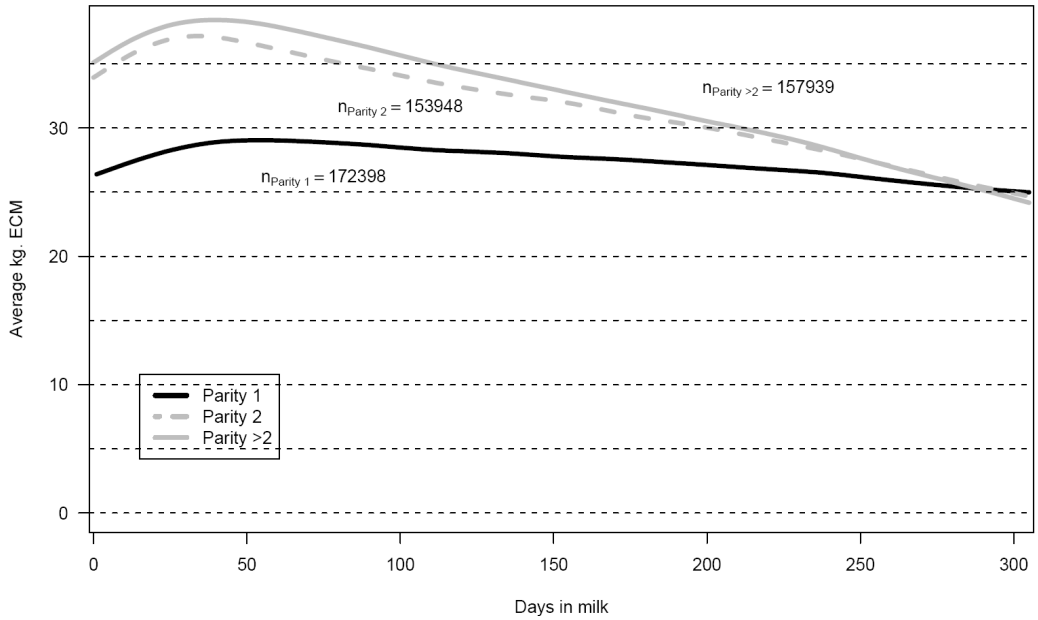


Figure 1. Parity specific kg energy corrected milk (kg ECM) as a function of DIM for all cows in the 64 Danish dairy herds.

Table 1. Parity specific distribution of herd level averages of milk production and corrected optical density (OD_C) values

	Parity 1				Parity 2				Parity 3+			
	Mean	Q1	Q2	Q3	Mean	Q1	Q2	Q3	Mean	Q1	Q2	Q3
Energy-corrected milk (kg/d)	27.8	24.4	27.9	31.3	32.5	27.5	32.5	37.4	33.4	27.8	33.4	39.1
ELISA test-value (OD_C)	0.05	0.00	0.00	0.00	0.12	0.00	0.00	0.10	0.15	0.00	0.10	0.20

Q1, Q2, Q3 = quartiles 1, 2 and 3, respectively.

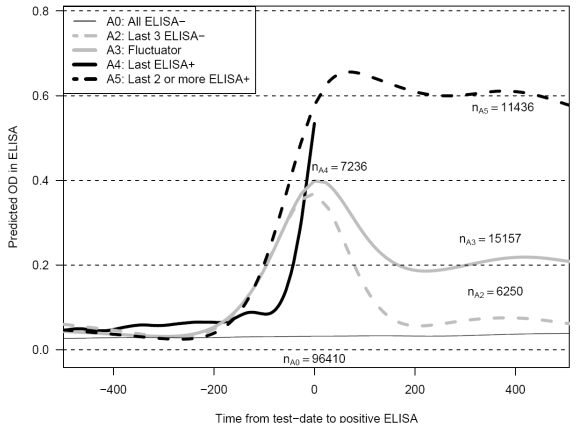


Figure 2. Average optical density (OD) value in ELISA for cows in the antibody groups relative to the date of first test-positive ELISA-test (antibody groups A2 to A5) or relative to the difference between first and last ELISA-negative test-result (antibody group A0).

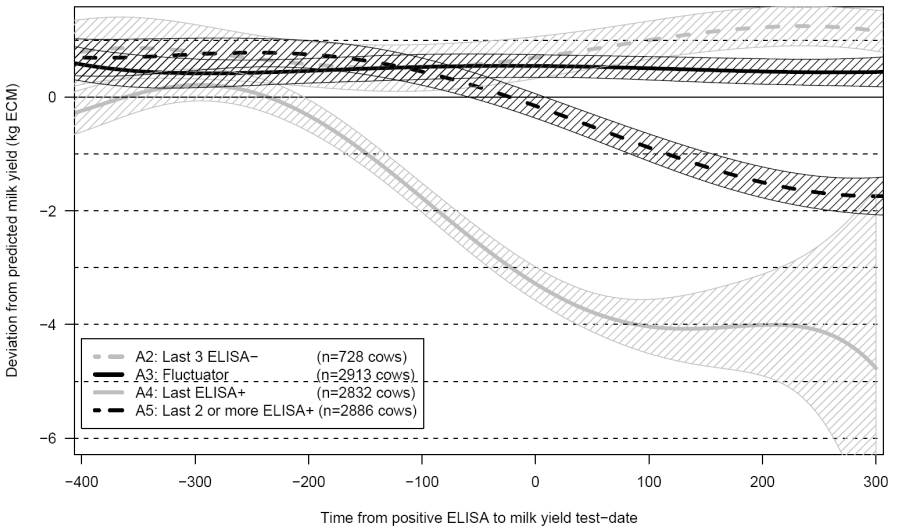


Figure 3. Deviation from predicted milk yield as a function of time from first positive ELISA test for cows from antibody groups A2 to A5, relative to the expected milk yield based on cows from antibody group A0 in the k^{th} herd in parity groups 1, 2 or >2. Shaded areas represent 95% confidence intervals for each data point.

The expected kilograms of ECM for cows in antibody group A0 were calculated and are represented by the value 0 kg of ECM in Figure 3. Deviations from 0 kg of ECM for cows with other antibody profiles are also shown in Figure 3. Mean milk yield of cows in antibody groups A2 and A3 were >0.5 kg of ECM greater per test date from 400 d before D 0 until 300 d after D 0. In addition, group A5 had a greater milk yield until 100 d before D 0. From –100 to +300 d relative to D 0, test-day milk yield decreased 2.3 kg of ECM in antibody group A5. Test-day milk yield decreased 4 kg of ECM in antibody group A4, beginning approximately 300 d before D 0 and with an apparent maximum decrease around 100 d after D 0.

Approximately 100 d after D 0, the uncertainty (visualized through the 95% confidence band) related to the decline in milk production for group A4 was much greater than previously, indicating that the number of observations from this period was reduced, probably because of culling or death of the animals. The uncertainty associated with the deviation in milk production for cows in antibody group A5 was smaller compared with the uncertainty of cows in antibody group A4 at 200 d after D 0.

DISCUSSION

The results showed that milk production was significantly reduced for cows when the last ELISA result was positive. The onset of this production loss occurred 300 d before the date the antibodies occurred, and the deviation from expected milk yield continued to increase after the cow had become antibody positive. Interestingly, cows that had repeated positive ELISA results had a less pronounced deviation in milk production loss compared with those with only 1 unconfirmed result. Two likely explanations are discussed further. It is likely that some cows cope better with the infection than others for a while, maybe because of a feed ration that supports the immune system of the cow (Stabel and Goff, 2004). Such a cow may be able to endure the infection for a longer time, but will eventually experience a decline in milk production. Another possible explanation is that the proportion of false-positive ELISA reactions was greater among cows that were repeatedly ELISA positive, thus diluting the effect caused by MAP infections. The latter explanation seems less likely, because the production loss occurred eventually, and the uncertainty associated with this loss was moderate.

The finding that all cows that become positive have a greater milk production 200 to 400 d before becoming test positive supports the findings of Wilson et al. (1993) and suggests that greater producing dairy cows were more likely to be test positive and eventually experience reduced milk production. These results would explain the discrepancies in the significance and magnitude of milk production losses described in previous studies (Johnson et al., 2001; Hendrick et al., 2005; Lombard et al., 2005). Our results provide support for the validity of our statistical approach, which allowed such patterns, in contrast to a model that attempts to estimate the contrasts between categories directly. We chose to use this model because the data would define the onset of the "disease process" (i.e., losses in milk production).

Culling based on 1 test-positive result is generally discouraged in the Danish control program, unless other information (such as a decline in milk production) is used in combination with the ELISA results (Nielsen et al., 2007). The results from the current study suggest that cows in antibody groups A4 and A5 could be potential candidates for culling based on their milk production, whereas cows in groups A2 and A3 were not, although, in a control plan on paratuberculosis, the risk of shedding of MAP from these cows should be

considered to reduce the transmission of MAP. Cows in antibody group A2 previously had a low probability of shedding MAP on a given day in the following year (2 to 3%), whereas the probability increased for cows in groups A3, A4, and A5 (Nielsen, 2008). Cows in groups A4 and A5 had high risks of shedding MAP, either on the test date or in the following test period. They may thus be contributing to the transmission of MAP, and culling decisions should include this aspect. Cows in group A3 may not be obvious candidates for culling from a milk loss point of view. Therefore, an alternative to culling could be to ascertain that the cow does not transmit MAP to susceptible cattle in the herd. A potential negative effect on the production economy caused by culling based on false-positive test reactions could be reduced.

A weakness of the current study was the use of a nonperfect test, which was neither 100% sensitive nor 100% specific. This is a general problem in studies on MAP. Confirmatory testing (e.g., using FC) is not used in the Danish program, although the possibility exists. Confirmatory testing could have reduced the number of false positives, but the number of false negatives would have increased. From a practical point of view, it was considered important to mimic the situation of the farmers. Because the production loss occurred before test positivity, and because a FC result can be achieved only 2 to 4 mo after the ELISA result, the usefulness of FC would be limited. In addition, an animal might be in a meager body condition if her milk production was already affected. Therefore, in practice a combination of deviation in milk production compared with the expected milk yield, combined with a test-positive result would be a much better (timely) basis for decisions on culling than an ELISA result combined with a confirmatory FC result.

A major strengths of this study were the availability of data from a huge population, which generally resulted in low uncertainty of the deviations in milk production, and the use of the predictive value of the individual cow in the individual herd. The variation that did exist was used to further explain the results. The results can be used directly by farmers to predict the fate of a cow, although better models may be developed when more test-day records become available. Nevertheless, the best (and least expensive) option to reduce the probability of uncertainty is the availability of more frequent ELISA tests per cow. The results can be used in the development of more precise models for estimating milk production losses on the herd and national level (including sires), because divisions into antibody groups are strong predictors for milk production losses. Another important finding was the long time span from the initial production loss to the maximum production loss. Such losses may not be realized by the herd manager without more advanced management tools, such as the model proposed here. Illustration of this chronicity problem has not been possible based on a sample representative of a major dairy population. It is important to note that cows with various antibody profiles experience different production losses, which is probably due to the capability of some cows to control the infection even when antibodies have been produced (Koets et al., 2001).

CONCLUSION

Major milk production losses for cows with certain well-defined paratuberculosis antibody profiles can commence 300 d before the first positive antibody test. Cows with fluctuating antibody profiles do not experience losses. Predictions of milk production loss from paratuberculosis are substantially more precise and useful when they are based on repeated test results compared with single test results.

REFERENCES

Benedictus, G., A. A. Dijkhuizen, and J. Stelwagen. 1987. Economic losses due to paratuberculosis in dairy cattle. Vet. Rec. 121:142-146.

Bennedsgaard, T. W., C. Enevoldsen, S. M. Thamsborg, and M. Vaarst. 2003. Effect of mastitis treatment and somatic cell counts on milk yield in Danish organic dairy cows. J. Dairy Sci. 86:3174-3183.

Hastie, T. J., and R. J. Tibshrinani. 1990. Generalized additive models. Chapman & Hall, London, England.

Hendrick, S., D. Kelton, K. Leslie, K. Lissemore, M. Archambault, and T. Duffield. 2005. Effect of paratuberculosis on culling, milk production and milk quality in dairy herds. J. Am. Vet. Med. Assoc. 227:1302-1308.

Johnson, Y. J., J. B. Kaneene, J. C. Gardiner, J. W. Lloyd, D. J. Sprecher, and P. H. Coe. 2001. The effect of subclinical *Mycobacterium paratuberculosis* infection on milk production in Michigan dairy cows. J. Dairy Sci. 84:2188-2194.

Kennedy, D. J., and S. S. Nielsen (eds.). 2007. Report from the first IDF ParaTB Forum. Bulletin of the International Dairy Federation. 410:3-7.

Koets, A. P., V. P. M. G. Rutten, M. de Boer, D. Bakker, P. Valentin-Weigand, and W. van Eden. 2001. Differential changes in heat shock protein-, lipoarabinomannan-, and purified protein derivative-specific bovine *Mycobacterium avium* subsp. *paratuberculosis* infection. Infect. Immun. 69:1492-1498.

Lombard, J. E., F. B. Garry, B. J. McCluskey, and B. A. Wagner. 2005. Risk of removal and effects on milk production associated with paratuberculosis status in dairy cows. J. Am. Vet. Med. Assoc. 227:1975-1981.

Nielsen, S. S. 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subsp. *paratuberculosis* in dairy cattle. J. Vet. Med. B 49:384-387.

Nielsen, S. S. 2008. Transitions in diagnostic tests used for detection of *Mycobacterium avium* subsp. *paratuberculosis* infections in cattle. Vet. Microbiol. 132, 274-282.

Nielsen, S. S., Ø. R. Jepsen, and K. Aagaard. 2007. Control programme for paratuberculosis in Denmark. Bulletin IDF-FIL. 410:23-29.

Sergeant, E. S., S. S. Nielsen, and N. Toft. 2008. Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds. Prev. Vet. Med. 85:92-106.

Stabel, J. R., and J. P. Goff. 2004. Efficacy of immunologic assays for the detection of Johne's disease in dairy cows fed additional energy during the periparturient period. J. Vet. Diagn. Invest. 16:412–420.

Wilson, D. J., C. Rossiter, H. R. Han, and P. M. Sears. 1993. Association of *Mycobacterium paratuberculosis* infection with reduced mastitis, but decreased milk production and increased cull rate in clinically normal dairy cows. Am. J. Vet. Res. 54:1851-1857.

Paper VII

A review of prevalences of paratuberculosis in farmed animals in Europe

Nielsen SS, Toft N

Published in
Preventive Veterinary Medicine 2009, 88, 1–19

A review of prevalences of paratuberculosis in farmed animals in Europe

Søren Saxmose Nielsen[a] and Nils Toft[b]

[a]Department of Large Animal Sciences, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark; [b]Danish Meat Association, Vinkelvej 11, DK-8620 Kjellerup, Denmark

ABSTRACT

Prevalence estimates are used by decision makers such as policy makers and risk assessors to make choices related to certain diseases and infections. Paratuberculosis, caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP), is a chronic infection particularly resulting in economic losses among farmed ruminants. Therefore, this infection is of concern for many farming industry decision makers. As a result, multiple studies have been carried out to determine the within-herd and between-herd prevalence of MAP infections. The objective of the present study was to conduct a review and, if possible, compare animal and herd level prevalences of MAP among farmed animals in Europe.

European data on prevalences of MAP in all farmed animal species were included from a review of literature. Information on target population and study design, tests used and apparent prevalences were recorded, and subsequently true prevalences were calculated when possible. A critical review of the included studies indicated that although a wide range of studies have been conducted, likely and comparable true prevalence estimates could rarely be calculated. Based on a few studies where the prevalences appeared to be plausible, it was concluded that prevalences of MAP would have to be guesstimates based on available data. The true prevalence among cattle appeared to be approximately 20% and was at least 3–5% in several countries. Between-herd prevalence guesstimates appeared to be >50%. No countries had published sufficient information to claim freedom from MAP or just a near-zero prevalence of MAP infections. No within-flock prevalence estimates were available for goats and sheep. The between-flock prevalence guesstimates were >20%, based only on estimates from Switzerland and Spain.

Major critical issues were identified in the majority of the studies, primarily due to lack of knowledge of the test accuracy of the diagnostic test used, or due to studies where the study population did not reflect the target population. Because most of the reported studies did not yield prevalence estimates which were comparable to other studies, there is a continuing need for well-designed studies of the prevalence of MAP infections.

1. INTRODUCTION

Paratuberculosis is a chronic infection of ruminants and other animals, caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). Primary industries have a particular interest in the infection and programmes have been established to control MAP infections (e.g. Sternberg and Viske, 2003; Franken, 2005; Carter, 2007; Kennedy, 2007; Kennedy and Benedictus, 2001; Kobayashi et al., 2007; Kulkas, 2007; Nielsen et al., 2007a). Major reasons for intervention against MAP are economic losses experienced by farmers (Ott et al., 1999) as well as the possible role of MAP in Crohn's disease in humans (Anon, 2000).

The prevalence of an infection at the herd and animal level is often a key issue when decision or policy makers determine whether the infection should be considered important or

not, and which measures to apply. Such measures could pertain to eradication in case of a low prevalence, control in case of a high prevalence, and surveillance in case of the likely absence of the infection in question. Prevalence estimates are also often required as input parameters in epidemiological models used for simulation of spread of the infection, in risk assessments etc. However, the prevalence estimate obtained in a survey is affected by the accuracy of the diagnostic test used and for comparison across studies the apparent prevalence (AP) needs to be adjusted with the relevant test characteristics in order to estimate the true prevalence (TP). The latency and slow progression of MAP infection makes diagnosis a challenge and tests with high sensitivity (Se) and specificity (Sp) are lacking (Nielsen and Toft, 2008). Furthermore, there is often discrepancy between the case definition used in the test evaluation study and the case definition used in the prevalence study, resulting in biased estimates of Se and Sp.

The objectives of the present study were to conduct a critical, comparative review of current within- and between-herd prevalence estimates among farmed animals in Europe, and whenever possible, to estimate TP corrected for the test-characteristics of the tests used in the studies. Ultimately, the aim was to determine if there were studies which could be used to produce valid prevalence estimates of MAP in Europe.

2. MATERIALS AND METHODS

2.1 Review of prevalence studies in literature

The case definition or target condition for “paratuberculosis” used for reporting of prevalence estimates in this review regarded prevalences of animals infected with MAP, unless otherwise specifically stated. The target condition was animals carrying MAP intracellularly, but substantial replication needed not to have taken place, because the infection could be latent, i.e. we used the widest of the three target conditions defined in Nielsen and Toft (2008). For herds, the case definition of a MAP infected herd included premises (herds or flocks) in which infected farm animals were present, or premises where MAP were present in the environment without necessarily having infected animals.

A review of prevalence studies was carried out by searching the available databases by April 12, 2007. Only studies from 1990 to present were considered relevant. The search databases included: Agricola (January 1990 to March 2007); Agris (January 1989 to December 2006); Biological Abstracts/RPM (R) (January 1990 to December 2000); BIOSIS Previews (January 2004 to December 2006); Biological Abstracts (January 1990 to December 2000), CAB Abstracts (January 1989 to February 2007) through ERL® Webspirs, and Medline through Pubmed (January 1990 to April 12, 2007). The search terms were: paratuberculosis, Johne's or Johnes combined with prevalence, incidence or occurrence. The same search terms were applied to the proceedings of the 6^{th}, 7^{th} and 8^{th} International Colloquia on Paratuberculosis in 1999, 2002 and 2005, respectively. Studies from these proceedings were included if they had not appeared in peer reviewed journals.

2.2. Data extraction and quality assessment

Data extraction was conducted in several steps. First, publications not reporting prevalences and publications reporting information from one herd only were excluded. Also excluded were reports of prevalences from non-European countries, along with studies conducted prior to 1990 and studies where farmed animals were not in the study population, i.e. data from zoo and game parks were excluded. Subsequently, data were extracted from the

remaining publications in a standard form including the following items: (a) country and region; (b) year of study; (c) animal species; (d) test used for classification of animals; (e) age distribution of study population; (f) number of animals and herds studied; (g) number of animals and herds test-positive; and (h) sampling strategy. The quality of the studies was generally poor, and many important details were only vaguely described or missing, e.g. an unclear definition and description of the target and study populations, or lack of reporting of period of study. The only criteria used for the subsequent inclusion was therefore the reporting of: a prevalence, country, animal species, test used, no. of animals or herd studied, no. of animals or herds test-positive and sampling strategy. In all studies included, vaccinated herds/flocks were excluded from the estimates reported. The reason for this exclusion was the potential long-lasting effect on the immune-response detected by serological tests, which can be expected following vaccination (Muskens et al., 2002).

Rather than excluding studies with critical issues, notes describing the relevant important concerns were attached to these studies. Critical issues in estimation and interpretation of TP are the test accuracy (i.e. Se and Sp) and the sampling frame used for obtaining samples from the target population. Non-random sampling leads to selection bias, e.g. “healthy worker survivor bias” (Arrighi and Hertz-Picciotto, 1994), which can result if the oldest animals are selected. Comparison of prevalences is dependent on the calculation of TPs, if tests with different accuracies are used. Test accuracies are often estimated based on non-random sampling, and often not estimated to address the target condition in question (Nielsen and Toft, 2008). If no valid test accuracy estimates exist for a test used in a prevalence study, extrapolation from estimates obtained for other tests for the same target condition may be used. However, unlikely or uncertain TPs may then result, partly because the test accuracy may only be valid for the target population in which a test has been validated (Greiner and Gardner, 2000), partly because one target condition have been used for evaluation of the test (usually with animals shedding MAP as a reference category, i.e. target condition is “infectious”) and another for prevalence estimation, where the target condition is usually “MAP infected”.

For animal level studies, the critical notes made in this review were: (i) the sampling frame used in the study design was not suitable for inference making in the target population; (ii) unlikely TPs were calculated based on the reported AP obtained with a given test with the most likely test accuracy (Table 1); (iii) prevalence estimates from the same country or region were obviously different (i.e. extremely high and extremely low); (iv) TP estimates provided in the original study were different than those which could be estimated with most likely test accuracy estimates, because the test accuracy estimates used originally were based on a different target-condition; (v) TP estimates could not be calculated because no test accuracy existed for the test used.

For herd level studies, critical notes included: (a) critical issues from animal level, which affected the interpretation of test results when aggregated on herd level; (b) a study design based on non-random sampling, e.g. inclusion of herds based on suspicion or selection of the oldest animals in a herd; (c) a non-specified sampling frame where it could not be verified if random sampling had been used; (d) assessment of a target condition different from “MAP infected”. The critical notes at both animal and herd level were used to assess if a prevalence estimate deduced from a given study could be compared to prevalence estimates from other studies.

Table 1. Test sensitivities and specificities used for calculation of true prevalences based on apparent prevalences among farmed animals. Estimates are medians of those obtained in a review of sensitivities and specificities among infected animals for a given test with a given antigen (summarised from Nielsen and Toft (2008) based on the tests used in the prevalence studies reported here)

Species	Test	Test antigen†	$Se_{Summarised}$	$Sp_{Summarised}$
Cattle	Faecal culture	NA	0.25	0.980
	Milk ELISA	PPA3	0.11§	0.985§
	Serum ELISA	PPA3	0.11#	0.985#
		VRI316	0.146	0.990
		IDEXX	0.088	0.976
		LAM	0.182	0.935
		Pourquier	0.15	1.0
		Various	0.15¤	0.980¤
Goat	Faecal culture	NA	0.08	0.980
	Serum ELISA	IDEXX	0.730	0.975
		Pourquier	0.778	1.00
		VRI316	0.842	0.997
Sheep	Faecal culture	NA	0.16	0.970
	Serum ELISA	IDEXX	0.37	0.985
		VRI316	0.317	0.982
Sheep & Goats combined			Cannot be estimated, as the test accuracy appear to differ between species	
Other species	No studies have been performed			

§) Derived from Nielsen et al., 2007b, because no other were available

#) According to Nielsen et al., 2002 and Sweeney et al., 1994, there is no difference in the use of milk and serum ELISAs. Therefore, the same estimates as the milk ELISA were used, because no separate estimates existed for the serum ELISA

¤) Based on medians of sensitivity and specificity of other ELISAs

†)Test antigens were not specified for the Pourquier and IDEXX tests, and are therefore simply given the test names.

2.3. Data analysis

The AP on animal and herd level was calculated as the number of test-positive animals among the total number of animals tested, or test-positive herds among the total number of herds tested (AP = T+/n). On animal level, test-positive was defined by the recommendation of those performing the study. On herd level, the cut-off for definition of a positive herd was 1 test-positive animal in case an animal level test was used, irrespective that in some studies, the test used was less than 100% specific. This approach was used in most publications, although results at multiple cut-offs were reported in some studies.

An estimate of the true prevalence (TP) can be obtained from the AP, if the test accuracy is known, by correction via the Rogan-Gladen estimator (Rogan and Gladen, 1978):

$$AP = Se \times TP + (1 - Sp) \times (1 - TP) \Leftrightarrow TP = \frac{AP + Sp - 1}{Sp + Se - 1}$$

The herd level counterparts of Se, Sp, AP and TP can be termed HSe, HSp, HAP and HTP, respectively. There is little difference between the animal level and herd level parameters, however, if an imperfect animal level test has been used as an aggregate test to classify a herd, the HSe can be calculated as: $HSe = 1 - (1 - AP)^n$, where n is the number of

animals tested in the herd. The HSp can be calculated as $HSp = Sp^n$ (Martin et al., 1992; Christensen and Gardner, 2000). As with the animal level conversion between AP and TP, the conversion between HAP and HTP requires that HSe and HSp is known.

To calculate comparable TPs based on the reported APs, we used median estimates of Se and Sp (Table 1) for a given test and test antigen, derived from the review in Nielsen and Toft (2008), as well as TPs calculated based on the diagnostic test accuracies reported in the original studies if any were available.

3. RESULTS

3.1 Literature search

The initial search generated 1189 hits (810 obtained through ERL® Webspirs and 379 through PubMed) including duplicate records. After exclusion of duplicate records, 701 publications remained. Another 34 were excluded because they were published before 1990.

The abstracts of the remaining 667 publications were evaluated and further reduction was performed excluding studies from non-European countries, studies on non-farmed animals, non-prevalence studies and studies including only one herd. The remaining 34 publications comprised prevalence studies as summarised in Table 2, including studies from Austria, Belgium, Czech Republic, Denmark, France, Germany, Greece, Italy, Norway, Portugal, Rep. of Croatia, Rep. of Ireland, Slovenia, Spain, Sweden, Switzerland, The Netherlands, Turkey and United Kingdom. These 34 publications also included 2 prevalence studies which were obtained by reviewing the proceedings of the 6th, 7th and 8th International Colloquia on Paratuberculosis.

Table 2. Distribution of the number of studies on prevalence of paratuberculosis in farmed animals in Europe

Family	Species	No. of studies	
		Animal level	Herd level
Artiodactyls, Bovidae	Cattle	30	23
	Goats	5	4
	Sheep	6	4
	Mixed goats and sheep	2	2
Artiodactyls, Camelidae	Llama	1	0
Artiodactyls, Cervidae	Fallow deer	1	0
	Red deer	1	0

Among the 34 publications reporting on prevalences in farmed animals, 46 were animal level studies and 33 were herd level studies. We used the term ‘study’ to define a study pertaining to one population of a specific animal species tested with one test. For example, some reports included both animal level and herd level estimates, and these each counted as one study. Reports where two tests were used on the same population also counted as two studies.

3.2. Cattle

The reported APs of MAP in cattle in various European countries are summarised in Table 3. Also included is information regarding the test used, sample sizes, study population and additional comments of importance for the calculation and interpretation of the TPs. TPs in Table 4 were calculated based on the test accuracy estimates given in Table 1 and APs given in Table 3 for each study, where test accuracy estimates was available. HAP estimates of MAP infections are summarised in Table 5, including information similar to that on cow

level. In the following sections, critical comments are given to the resulting TPs, and the potential credibility of the prevalence estimates obtained.

Table 3. Animal level apparent prevalences of infection with *Mycobacterium avium* subsp. *paratuberculosis* in cattle in Europe, reported in literature January 1990 to April 12, 2007

Country / Region	Study period	Test	Age-group	n	T+	AP (%)	Reference[#]
Austria /all	?	ELISA, Svanovir	Cull cattle > 2 yrs	756	144	19.0	Dreier et al., 2006[a]
Austria / 4 regions	1995-97	ELISA, Allied	Four oldest /herd	11028	664	6.0	Gasteiner et al., 1999
Belgium /all	1997-98	HerdChek, IDEXX, France	> 2 yrs	13150	116	0.9	Boelaert et al., 2000
France /Yonne	1998-99	Pourquier ELISA	All	8793	292	3.3	Petit, 2001
Germany /Saxony	2002-2004	HerdChek, IDEXX Wörrstadt, Germany	?	3454	151	4.4	Donat et al., 2005[b]
Germany /Arnsberg	1993	ELISA, in-house	Cull cattle > 18 mo.	536	79	15	Böttcher, 1997
Germany /Arnsberg	1993	Tissue culture	Cull cattle > 18 mo.	517	7	1.4	Böttcher, 1997
Germany /Bavaria	2005	HerdChek, IDEXX, Wörrstadt, Germany		2748	41	1.5	Böttcher and Gangl, 2004
Germany /Bavaria	2005	Pourquier ELISA		2748	25	0.9	Böttcher and Gangl, 2004
Germany /Bavaria	2005	ELISA, Svanovir		2748	662	24	Böttcher and Gangl, 2004
Italy /Latium	?	HerdChek, IDEXX		19627	472	2.4	Lillini et al., 2005,
Italy /Veneto	2000-2001	HerdChek, IDEXX	> 12 months	27135	949	3.5	Robbi et al., 2002
Italy /Verona province	1997-98	Tissue culture + histopathology	?	73	5	6.8	Vicenzoni et al., 1999
Norway /?	1996-97	HerdChek, IDEXX	?	9456	728	7.7	Tharaldsen et al, 2003[c]
Norway /all	2002	Faecal culture	5 oldest in herd	1592	2	0.1	Mørk et al., 2003[d]
Rep. of Ireland /imported animals	1997	Parachek ELISA	?	225	8	3.6	O'Doherty et al., 2002[e]
Slovenia /all	1997	HerdChek, IDEXX, USA		11513	47	0.4	Ocepek et al., 1999
Slovenia /all	1998	HerdChek, IDEXX, USA		12082	140	1.2	Ocepek et al., 1999
Slovenia /all	2000-01	Pourquier ELISA	> 2 yrs	9388	41	0.4	Ocepek et al., 2002
Slovenia /all	1999	HerdChek, IDEXX, Sweden		38469	1305	3.4	Ocepek et al., 2002
Sweden /all	2000-2001	Faecal culture	?	4000	0	0.0	Sternberg and Viske, 2003[f]
Sweden /all	1995-96	Tissue culture	> 2 years	3166	1	0.0	Viske et al., 1996
Switzerland /Plateau de Diesse	?	Parachek ELISA, CSL, Australia	> 18 months	565	29	5.1	Meylan et al., 1995[g]
Switzerland /all	1993-94	Parachek ELISA, CSL, Australia	Dairy	1663	12	0.7	Stärk et al., 1997
Switzerland /?	2005	MAP F57 PCR	Cull cattle	101	20	19.8	Bosshard et al., 2006
The Netherlands /all	1998	HerdChek, IDEXX, Scandinavia	Dairy ≥ 3 years	15745	400	2.5	Muskens et al., 2000
Turkey /?		ELISA in-house	> 2 yrs	8873	409	4.6	Atala and Akcay, 2001
Turkey /Trakya		IS*900* PCR in faecal samples		96	0	0.0	Ikiz et al., 2005
Turkey /Elazig	1997-98	IS*900* PCR in milk	Dairy > 2 years	500	25	5.0	Cetinkaya et al., 2000[h]
United Kingdom /South west England	1994	IS*900* PCR on lymph nodes	Cull cattle	1297	46	3.5	Cetinkaya et al., 1996
United Kingdom /South west England	1994	Tissue culture of T+ from IS*900* PCR	Cull cattle	1297	34	2.6	Cetinkaya et al., 1996

Abbreviations: AP = apparent prevalence ; FC = faecal culture; n = number tested; T+ = test positive;

[#]) Comments: a) Only 4 /756 were FC positive which questions the ELISA or the FC; b) Many herds were selected based on suspicions. These were excluded; c) Sampling scheme unclear; d) The numbers reported do not match the surveillance scheme; e) Imported animals only; Only summary of survey. No details provided; g) Apparently not a random sample; h) Sampling frame unknown; i) Screening of tissues with IS*900* PCR with confirmation of T+ by tissue culture

Table 4. Animal level true prevalences of infections with *Mycobacterium avium* subsp. *paratuberculosis* in cattle in Europe based on studies reported in literature 1990 to April 12, 2007

Country / Region	Test	Antigen	AP (%)	Se	Sp	TP (%)[#]	Reference	Critical issues[§]
Austria /all	ELISA, Svanovir	LAM	19.0	0.182	0.935	107	Dreier et al., 2006	a
Austria /4 regions	ELISA, Allied	PPA3	6.0	0.11	0.985	47	Gasteiner et al., 1999	a
Belgium /all	HerdChek, IDEXX, France	IDEXX	0.9	0.088	0.976	-23	Boelaert et al., 2000	d
France /Yonne	Pourquier ELISA	Pourquier	3.3	0.15	1.0	22	Petit, 2001	
Germany /Saxony	HerdChek, IDEXX, Germany	IDEXX	4.4	0.088	0.976	31	Donat et al., 2005	
Germany /Arnsberg	ELISA, in-house	Various	14.7	0.15	0.98	98	Böttcher, 1997	b, e
Germany /Arnsberg	Tissue culture	Various	1.3				Böttcher, 1997	b, e
Germany /Bavaria	HerdChek IDEXX, Germany	IDEXX	1.5	0.088	0.976	-14	Böttcher and Gangl, 2004	b
Germany /Bavaria	Pourquier ELISA	Pourquier	0.9	0.15	1.0	6	Böttcher and Gangl, 2004	b
Germany /Bavaria	ELISA, Svanovir	LAM	24.1	0.182	0.935	150	Böttcher and Gangl, 2004	a, b
Italy /Latium	HerdChek, IDEXX	IDEXX	2.4	0.088	0.976	0.0	Lillini et al., 2005	b
Italy /Veneto	HerdChek, IDEXX	IDEXX	3.5	0.088	0.976	17	Robbi et al., 2002	
Italy /Verona province	Tissue culture and histopathology	NA	6.8				Vicenzoni et al., 1999	
Norway /?	HerdChek, IDEXX	IDEXX	7.7	0.088	0.976	83	Tharaldsen et al, 2003	b, c
Norway /all	Faecal culture	NA	0.1	0.25	0.98	-8	Mørk et al., 2003	b, c
Rep. of Ireland /imported animals	Parachek ELISA, CSL, Australia	VRI316	3.6	0.146	0.990	19	O'Doherty et al., 2002	c
Slovenia /all	HerdChek, IDEXX, USA	IDEXX	0.4	0.088	0.976	-31	Ocepek et al., 1999	b, d
Slovenia /all	HerdChek, IDEXX, USA	IDEXX	1.2	0.088	0.976	-19	Ocepek et al., 1999	b, d
Slovenia /all	Pourquier ELISA	Pourquier	0.4	0.15	1.0	3	Ocepek et al., 2002	b, d
Slovenia /all	HerdChek, IDEXX, Sweden	IDEXX	3.4	0.088	0.976	16	Ocepek et al., 2002	b, d
Sweden /all	Faecal culture	NA	0.0	0.25	0.98	-9	Sternberg and Viske, 2003	c, d
Sweden /all	Tissue culture	NA	0.0				Viske et al., 1996	c, d
Switzerland /Plateau de Diesse	Parachek ELISA, CSL, Australia	VRI316	5.1	0.146	0.990	30	Meylan et al., 1995	b, c, d
Switzerland /all	Parachek ELISA, CSL, Australia	VRI316	0.7	0.146	0.990	-2	Stärk et al., 1997	b, d
Switzerland /?	MAP F57 PCR	NA	19.8				Bosshard et al., 2006	b ,d
The Netherlands /all	HerdChek, IDEXX, Sweden	IDEXX	2.5	0.088	0.976	2	Muskens et al., 2000	d
Turkey /?	ELISA various	Various	4.6	0.15	0.98	20	Atala and Akcay, 2001	
Turkey /Thrace	IS*900* PCR in faecal samples	NA	0.0				Ikiz et al., 2005	c, d
Turkey /Elazig	IS*900* PCR in milk	NA	5.0				Cetinkaya et al., 2000	c, d
United Kingdom /South west England	IS*900* PCR on lymph nodes	NA	3.5				Cetinkaya et al., 1996	
United Kingdom /South west England	Tissue culture of T+ from IS*900* PCR	NA	2.6				Cetinkaya et al., 1996	

[#]) TP should be between 0 and 1, but was in some instances been calculated to <0 or >1. In those cases, the TP should be 0 and 1, respectively.
[§]) Critical issues identified: a) The test and test accuracy used produced incredibly high TPs; b) discrepancies from comparison with other estimates from same country or region; c) the sampling frame used in the study design was not suitable for inference making in the target population; d) TP estimates not interpretable; e) discrepancies between calculated TPs provided originally compared to the TP calculated with most likely test accuracy estimates. Further explanations are given in the text. Abbreviations: AP=apparent prevalence; Se=sensitivity; Sp=Specificity; TP=True Prevalence; NA=not applicable

Table 5. Summary of herd level apparent prevalences of *Mycobacterium avium* subsp. *paratuberculosis* infections in cattle in Europe, published from January 1990 to April 12, 2007.

Country /Region	Year	Test	Age	n	T+	HAP (%)	Critical issues§	Reference[#]
Belgium/ all	1997-98	HerdChek, IDEXX, France	> 2 yrs	458	82	18	A, D	Boelaert et al., 2000
Denmark /all	1998	Bulk tank milk ELISA	NA	900	497	55	D	Nielsen et al., 2000[a]
France /Yonne	1998-99	Pourquier ELISA	All	155	105	68		Petit, 2001
Germany /Saxony	2002-2004	HerdChek, IDEXX, Wörrstadt, Germany	?	493	207	42	B	Donat et al., 2005[b]
Italy /Latium	?	HerdChek, IDEXX		369	155	42	A	Lillini et al., 2005
Italy /Veneto	2000-2001	HerdChek, IDEXX	> 12 months	419	272	65		Robbi et al., 2002
Italy /Verona	1997-98	Tissue culture and histopathology	?	73	5	7		Vicenzoni et al., 1999[c]
Norway /All	2002	Faecal culture	5 oldest in herd	200	2	10	B	Mørk et al., 2003[d]
Rep. of Ireland /all	2000-01	Bulk tank milk IS*900* PCR	NA	389	50	13	D	O'Reilly et al., 2004
Slovenia /all	1997	HerdChek, IDEXX, USA		1690	48	3	A	Ocepek et al., 1999
Slovenia /all	1998	HerdChek, IDEXX, USA		2423	157	7	A	Ocepek et al., 1999
Slovenia /all	2000-01	Pourquier ELISA	> 2 yrs	302	35	12	A	Ocepek et al., 2002
Slovenia /all	1999	HerdChek, IDEXX, Sweden		26088	1055	4	A	Ocepek et al., 2002
Spain /all	?	Bulk tank milk IS*900* PCR	Lactating cows	70	7	10	D	Sevilla et al., 2002
Spain /Northern Spain	?	Bulk tank milk IS*900* PCR	Lactating cows	200	16	8	D	Sevilla et al., 2002
Sweden /all	2000-2001	Faecal culture	?	200	0	0	A, C	Sternberg and Viske, 2003[e]
Switzerland / North-east, central and north-west	2000-2001	Bulk tank milk IS*900* PCR	Cows	501	112	22	D	Stephan et al, 2002[f]
Switzerland /all	1993-94	Parachek ELISA	?	113	9	8	A	Stärk et al., 1997
Switzerland /all	2001	Bulk tank milk IS*900* PCR	Cows	1384	273	20	D	Corti and Stephan, 2002
Switzerland /?	2005	Bulk tank milk F57 PCR		100	3	3	D	Bosshard et al., 2006[h]
The Netherlands /all	1998	HerdChek, IDEXX, Scandinavia	≥ 3 years	371	200	54	A	Muskens et al., 2000
Turkey /?		ELISA, in-house	> 2 yrs	545	206	38		Atala and Akcay, 2001
United Kingdom / England and border regions of Wales	1995	Observations by farmers in 1993 & 94	?	2801	486	17	D	Cetinkaya et al., 1998[i]

#) Comments: a) A stratified random sample; b) Many herds were selected based on suspicions. On cow-level, these were excluded; c) Sampling frame not stated; d) The numbers reported are different from those in the surveillance scheme; e) Only summary of survey with no details; f) Unclear if these results were part of the study reported by Corti and Stephan (2002); g) Only recorded as ELISA positive if "++"; h) Selection of sample material unclear; i) Only clinical disease.

§) Critical issues: A) critical issues from animal level affected the interpretation of test-results when aggregated on herd-level; B) study design was based on non-random sampling, e.g. inclusion of herds based on suspicion or selection of the oldest animals in a herd; C) sampling frame not specified and it could therefore not be verified that random sampling had been used; D) a target condition different from "MAP infected" was assessed.

Abbreviations: n = number of herds test; T+ = number of test positive herds; HAP = herd level apparent prevalence; NA = Not applicable

3.2.1. Critical review of estimated true prevalences in cattle on animal level

Most animal level prevalence studies were given critical comments, which are summarised in Table 4. Eight studies were designed, so the samples were not representative of the population in the country or region, from which the study was reported. Imported animals were a primary target group in Republic of Ireland (O'Doherty et al., 2002) and Sweden (Viske et al., 1996; Sternberg and Viske, 2003), and the study components including imported animals were therefore not suitable for inference on the Irish and Swedish cattle populations. The number of cattle imported to Sweden was stated to be low, but these animals were primarily those subject to testing. Sternberg and Viske (2003) and Viske et al. (1996) also reported summaries of other studies, but few details were available on the design of those studies. Another main criticism that can be directed towards the design of some reported studies was the testing of the oldest animals only. These animals may have been old because they were non-infected. The five oldest animals were tested in each of 200 randomly selected Norwegian herds (Mørk et al., 2003), and in another study of the Norwegian cattle population (Tharaldsen et al., 2003), random selection was used for some animals only, and no separate estimates were given for those animals that were actually sampled randomly. Neither of the Norwegian studies therefore seemed representative of the Norwegian cattle population. Meylan et al. (1995) did not include a random sample of herds while studying prevalence of MAP in Switzerland, and in two prevalence studies from Turkey (Cetinkaya et al., 2000; Ikiz et al., 2005), the sampling frames were not specified.

We grouped issues related to test accuracies and comparison of within-study TPs together, because these issues are linked. From Sweden, Viske et al. (1996) reported the study of a random sample of adult (>2 years) cattle sent for slaughter. The diagnostic method used was culture of a pool of tissue from ileum and an ileocecal lymph node. Viske et al. (1996) assumed that the sensitivity was between 0.5 and 1.0. However, according to Whitlock et al. (1996) it is insufficient to sample tissues only from ileum and the ileocecal lymph nodes, because this sampling procedure will fail to detect MAP in many infected animals. Under these circumstances and with only one culture positive animal, it cannot be determined if the prevalence was low or the diagnostic test had low sensitivity.

In the study by Böttcher (1997) based on serum and tissues samples from animals sent to slaughter, APs were 14.7% based on an in-house ELISA and 1.4% based on tissue culture (Table 3). A Se of the ELISA of 0.44 and a Sp of 0.86 based on the tissue culture results was reported from the study. These original estimates of Se and Sp of the ELISA would result in a TP of 2.3%, and not in a TP of 98% as estimated using the Se and Sp estimates given in Table 1. The TP estimated by Böttcher (1997) seemed more reasonable given the results of tissue culture, although the discrepancy between the results render interpretation questionable.

Two studies (Böttcher and Gangl, 2004; Dreier et al., 2006) were based on the Svanovir test (Svanova Biotech AB, Uppsala, Sweden) and resulted in calculated TPs >100% (100% subsequent to rounding), which appears unlikely. In the report by Böttcher and Gangl (2004), estimates of prevalences based on two other tests (IDEXX HerdChek, Idexx Gmbh, Wörrstadt, Germany, and Pourquier ELISA, Institut Pourquier, Montpellier, France) yielded TPs of 0% and 6%, respectively, indicating that the TP of 100% was an invalid result. The TP of 100% was unlikely, because some animals would be expected not to be infected, particularly as some herds must be expected to be free of MAP. Böttcher and Gangl (2004) did not conclude if 0%, 6% or 100% was the more likely prevalence. No estimate gave the

impression of being more reasonable than others when compared, and interpretation did not seem possible.

Gasteiner et al. (1999) studied the prevalence of MAP in the regions Mödling, Graz, Linz and Innsbruck in Austria. They used an ELISA based on an antigen from Allied Monitor (Fayette, USA) and estimated the Se and Sp of this ELISA to 0.882 and 0.816, respectively. These test accuracy estimates would result in a calculated TP of 0%, while we estimated the TP to 47% (Table 4). The Se estimate given in the original study appeared unlikely, and the test evaluation was not performed to meet the target condition MAP infected animals. Our TP estimate could also be biased as the Se and Sp used were not necessarily valid for the ELISA used in the Austrian study. Therefore, the TP could not be determined for that study. In a study from Lazio, Italy, the TP based on estimates of Se and Sp given in Table 1 and those stated by Lillini et al. (2005) (Se = 0.77 and Sp = 0.95) would result in a TP of 0%. However, follow-up diagnostic testing using faecal culture, Ziehl-Neelsen staining of faecal samples and PCR was performed on a subset of 40 ELISA-positive herds and indicated that 31 of the herds were infected. Therefore, the TP could not be 0%. In other parts of Italy, the TP was estimated to 17% in the Venetia region based on an AP of 3.5% using ELISA (Robbi et al., 2002), which corresponded with the AP of 6.8%, based on histopathology and culture of MAP from tissue samples of the ileocecal valve, from animals sent to slaughter in the Verona province of Venetia (Vicenzoni et al., 1999). Although there were no Se and Sp estimates for tissue culture and histopathology, it would seem reasonably to infer that the TP was approximately in the range 10–20%, assuming that the tests were completely specific, but lacked sensitivity (Whitlock et al., 1996).

In Belgium, a stratified random sample of cows was obtained from 511 herds. In these herds, all cattle >24 months of age were sampled. The animals were tested by ELISA (HerdChek, IDEXX, France), and Boelaert et al. (2000) assumed that this test had a Se of 0.45 and a Sp of 0.99 based on data from the literature. This Se estimate was invalid, because it was based on a sample of animals where the distribution of infected and infectious was likely to be different from the distribution in the Belgian population (and most other populations). The median within-herd TP estimated originally was 7%, which may have been underestimated by a factor 5, considering that the Se was more likely to be 0.088 and not 0.45 (Table 1 and Nielsen and Toft, 2008). However, the within-herd TP could also be calculated to 0% (Table 4). It was therefore not possible to derive a TP estimate from the Belgium population based on published studies. Also in Slovenia, the calculated TPs shown in Table 4 indicated some discrepancies. Based on screenings (Ocepek et al., 1999) of stratified random samples performed in 1997 and 1998 using the IDEXX HerdChek test (Westbrook, Maine, USA), the TP would be 0%. A change of test to another IDEXX test from Sweden resulted in a TP of 16% in 1999 (Ocepek et al., 2002). In 2000–2001 screenings (Ocepek et al., 2002), the Pourquier test was used, resulting in a TP of 3%. Ocepek et al. (2002) stated that clinical paratuberculosis occurred in the entire period. Therefore, the TPs of 0% were unlikely, but it was not possible to determine which of the other estimates that was the more credible.

The studies that have been carried out in Switzerland also yielded TPs, which were not similar. Among the two studies based on Parachek ELISA (CSL, Parkville, Victoria, Australia), APs were estimated to 5.1% (Meylan et al., 1995) and 0.7% (Stärk et al., 1997), which subsequent to conversion gave TPs of 30% and 0%, respectively. Meylan et al. (1995) did not include a random sample of herds. However, their estimate was closer to the AP of

20% based on samples of multiple specimens (milk, faeces and tissues) tested by a PCR-method detecting MAP F57 sequences reported by Bosshard et al. (2006). The 101 animals included in the latter study were obtained among cull-cattle sent to one slaughter-house from all over Switzerland, but insufficient information was available to determine whether this sample was representative for the overall Swiss cattle population. It was therefore difficult to draw inferences on the prevalence in Switzerland. It was likely that a significant proportion of the animals were infected.

One Dutch study has been published based on a region-stratified random sample using the IDEXX HerdChek test (IDEXX Scandinavia AB, Österbybruk, Sweden) (Muskens et al., 2000). The Se of the original HerdChek test has been estimated to 0.088 (see Table 1). Muskens et al. (2000) decreased the Se of the test by changing the cut-off. They assumed that the Se was 0.3 and 0.4, but it seems more likely that the Se was much less than 0.088. Due to the change in the cut-off without an evaluation of the Se at this cut-off, the TPs were non-interpretable.

In three Turkish regions: (1) central Anatolia, (2) eastern Anatolia, and (3) western Anatolia combined with eastern Thrace, prevalence studies were carried out based on ELISA (Atala and Akcay, 2001), PCR in milk samples (Cetinkaya et al., 2000) and PCR in faecal samples (Ikiz et al., 2005), respectively. Se and Sp have not been estimated for the PCR methods and TPs cannot be calculated. In the study by Atala and Akcay (2001), an AP of 4.6% could be converted to a TP of 20%, but the estimate could not be verified. The only apparently interpretable Turkish study was therefore the study by Atala and Akcay (2001).

Prevalence estimates from United Kingdom have only been published in one study in south west England (Cetinkaya et al., 1996). This study was based on cull cattle sent to three abattoirs, and the representativity of the animals included could not be assessed but seemed reasonable. Of 1297 adult cattle tested, 3.5% were positive by an IS*900* PCR used for samples of lymph nodes. PCR-positive lymph nodes were also cultured, and in 34 (2.6% of the total sample) MAP was isolated. The sensitivity of the test used by Cetinkaya et al. (1996) has not been reported. Assuming that the test was 100% specific but less than 100% sensitive like other tests for ante-mortem diagnosis (Nielsen and Toft, 2008), a likely estimate of the TP would be minimum 3–5%.

Petit (2001) estimated the AP to be 3.3% among a random selection of cattle in the Yonne region of France. Assuming a Se of 0.15 and a Sp of 1.0 of the ELISA used (Pourquier, Montpellier, France), a TP of 22% was calculated. A similar AP of 4.4% among 3454 randomly selected animals from Saxony was obtained using a different ELISA (HerdChek, IDEXX Gmbh, Wörrstadt, Germany) (Donat et al., 2005). Using the Se and Sp given in Table 1 for the HerChek test, a TP of 31% could be estimated. Donat et al. (2005) stated a Se of 0.555 and a Sp of 0.948, but these estimates were obtained based on a sample of cattle that were not representative of infected animals. Therefore, the test accuracy estimates given in Table 1 would be more reasonable to use.

The true prevalence among cattle appears to be approximately 20%, based on the studies by Petit (2001) in France, Donat et al. (2005) in Germany, Robbi et al. (2002) in Italy, and Atala and Akcay (2001) in Turkey. Studies based on detection of MAP directly or pathological changes suggested that the prevalence is at least 3–5% in several countries (Cetinkaya et al., 1996; Vicenzoni et al., 1999). These figures should all be interpreted with caution considering the general uncertainties listed above.

3.2.2. Critical review of estimated true prevalences in cattle on herd level

Critical to the evaluation of the HTP is the TP on animal level in a herd. Many studies have been conducted to estimate both the herd level and the animal level prevalence, but corrections for Se and Sp prior to classification of the herd were not included in the study design and data analyses. Boelaert et al. (2000) provided an approach on how to address this issue. Their approach was sensitive to use of invalid test accuracy estimates, and the Se and Sp estimates they used appeared to be incorrect, resulting in non-interpretable cow level estimates, and thereby also the non-interpretable herd level estimates. This is a common problem in many of the studies reported, and the following studies resulted in non-interpretable HTPs: Boelaert et al. (2000), Lillini et al. (2005), Mørk et al. (2003), (Ocepek et al., 1999) and (Ocepek et al., 2002), Sternberg and Viske (2003), Stärk et al. (1997), Muskens et al. (2000) (Table 5). Non-interpretable TPs also resulted due to inclusion of non-random samples (Donat et al., 2005) and no specification of sampling strategy (Vicenzoni et al., 1999; Bosshard et al., 2006).

The study by Cetinkaya et al. (1998) was based on a postal survey where farmers were asked to report whether clinical paratuberculosis had occurred in 1993 or 1994. The Se and Sp of the test "farmers' observation" were not reported and clinical disease was not the target of the present review. However, the true prevalence of MAP infections would be expected to be higher than the observed prevalence, as clinical disease is only considered to occur in a fraction of those animals that are actually infected.

Among the remaining studies, three were based on animal level ELISA results, resulting in HAPs of 68% in Yonne, France (Petit, 2001), 65% in Veneto, Italy (Robbi et al., 2002) and 38% in Turkey (Atala and Akcay, 2001). The HTP could have been estimated using the approach suggested by Boelaert et al. (2000) but the raw data were not available and no estimates of HTP were provided. The resulting HTP in the Yonne-population would thereby be a minimum of 68%, assuming that the Sp of the test used was 1.0 as has previously been estimated Petit (2001).

Nielsen et al. (2000) estimated the HTP to be 47% as an average among a stratified random sample in Denmark, by use of antibody ELISA on bulk tank milk samples. An average of a stratified random sample is not a valid measure, and differences between the within-stratum specific estimates were huge. Bulk tank milk samples have also been used in several other studies, but these were all subject to PCR, either using the IS*900* sequence or the F57 sequence. In Spain, the HAPs in Northern Spain and all of Spain were 10% and 8%, respectively (Sevilla et al., 2002). In the Republic of Ireland, 13% of the samples were found positive by bulk tank milk PCR (O'Reilly et al., 2004). In Switzerland, one or two studies have been carried out resulting in HAP estimates of 22% (Stephan et al., 2002) and 20% (Corti and Stephan, 2002), bulk tank milk PCR positive, respectively. It is unclear whether the latter study was a subset of the former study, or the studies were two independent studies. Test accuracy estimates for the bulk tank milk PCRs were not available, but it was likely that they were nearly 100% specific, but lacked sensitivity as not all infected animals shed MAP in milk.

In summary, there were very few studies which provided interpretable estimates of herd level prevalences of MAP in cattle herds in Europe. Among studies detecting MAP sequences in milk, the HAPs were in the range 8–22%. These could be speculated to be at least two to three times higher, if converted to HTPs. Among studies based on antibodies, the HAPs were in the range 38–68%. A guesstimate on the herd level true prevalence would

therefore be that the prevalences in many countries were above 50%, maybe higher. In some areas they may be lower, but the documentation is so far insufficient.

3.3. Goats and Sheep

The number of prevalence studies carried out among small ruminants was small. An overview of the studies is given in Table 6, with calculated TPs in Table 7. TPs could not be estimated in two populations with mixed populations of sheep and goats. Flock level prevalences are summarised in Table 8, and critical comments to the studies are given in the following sections.

Table 6. Animal level apparent prevalences of *Mycobacterium avium* subsp. *paratuberculosis* infections among goats and sheep in Europe from studies published in January 1990 to April 12, 2007

Species	Country /region	Year	Test	Age	N	T+	AP (%)	Reference[#]
Goats	Austria /all	?	Faecal culture and tissue culture	?	80	0	0.0	Khol et al., 2006[a]
Goats	Austria /all	?	HerdChek, IDEXX Scandinavia	?	80	0	0.0	Khol et al., 2006 [a]
Goats	Norway /all	2002	Faecal culture	10 oldest in herd	662	7	1.1	Mørk et al., 2003[b]
Goats	Rep. of Croatia /?	2001	HerdChek, IDEXX, USA	?	375	0	0.0	Cvetnic, 2002 [a]
	Portugal /Lisbon	?	ELISA, Parachek, CSL	?	2351	41	1.7	Mendes et al., 2004[c]
Mixed sheep and goats	Portugal /Lisbon	?	ELISA, Parachek, CSL	?	2351	41	1.7	Mendes et al., 2004 [c]
Mixed sheep and goats	Slovenia /all	2000-01	Pourquier ELISA		12578	440	3.5	Ocepek et al., 2002[d]
Sheep	Austria /all	?	HerdChek, IDEXX Scandinavia	?	169	1	0.6	Khol et al., 2006 [a]
Sheep	Austria /all	?	Tissue culture and faecal cultue	?	169	1	0.6	Khol et al., 2006 [a]
Sheep	Norway /all	2002	Faecal culture	10 oldest in herd	369	1	0.3	Mørk et al., 2003[b]
Sheep	Rep. of Croatia /?	2001	HerdChek, IDEXX, USA	?	356	2	0.6	Cvetnic, 2002 [a]

#) Comments: a) Sampling frame not specified; b) The numbers reported are different from those in the surveillance scheme; c) Larger flocks and one flock suspected of clinical disease selected; d) No distinction between sheep and goats

Abbreviations: n = number of goats or sheep tested; T+ = number of test-positive; AP = apparent prevalence

Table 7. Estimates of animal level true prevalences of *Mycobacterium avium* subsp. *paratuberculosis* infections among goats and sheep in Europe based on information published in January 1990 to April 12, 2007

Species	Country /region	Test	AP (%)	Se	Sp	TP (%)[#]	Reference
Goats	Austria /all	Faecal culture and tissue culture	0.0				Khol et al., 2006
Goats	Austria /all	HerdChek, IDEXX, Scandinavia	0.0	0.73	0.975	-3	Khol et al., 2006
Goats	Norway /all	Faecal culture	1.1	0.08	0.980	-15	Mørk et al., 2003
Goats	Rep. of Croatia /?	HerdChek, IDEXX, USA	0.0	0.778	1.00	0	Cvetnic, 2002
Mixed sheep and goats	Portugal /Lisbon	ELISA, Parachek, CSL	1.7	NE	NE		Mendes et al., 2004
Mixed sheep and goats	Slovenia /all	Pourquier ELISA	3.5	NE	NE		Ocepek et al., 2002
Sheep	Austria /all	HerdChek, IDEXX, Scandinavia	0.6	0.37	0.985	-3	Khol et al., 2006
Sheep	Austria /all	Tissue culture	0.6				Khol et al., 2006
Sheep	Norway /all	Faecal culture	0.3	0.16	0.970	-20	Mørk et al., 2003
Sheep	Rep. of Croatia /?	HerdChek, IDEXX, USA	0.6	0.37	0.985	-3	Cvetnic, 2002

[#)] TP should be between 0 and 1, but has in some instances been calculated to <0 or >1. In those cases, the TP should be 0 and 1, respectively.

Abbreviations: AP = Apparent prevalence; NE = Not estimated; Se = sensitivity; Sp = specificity; TP = true prevalence

Table 8. Herd level apparent prevalences of infections with *Mycobacterium avium* subsp. *paratuberculosis* infections in goats and sheep in Europe based on reports published in January 1990 to April 12, 2007

Species	Country	Year	Test	Age group	n	T+	HAP (%)	Comments	Reference
Goats	Greece /Northern Sporades Islands	?	AGID & faecal culture	?	38	0	0	Information could not be verified	Minas et al., 1994
Goats	Spain /Avila	1996	AGID	> 18 months	23	12	52	Cross-sectional study of herds > 25 animals older than 18 months	Reviriego et al., 2000
Goats	Switzerland /all	2002	Bulk tank milk IS*900* PCR	NA	344	79	23		Muehlherr et al., 2003
Goats	United Kingdom /England, Wales, Northern Ireland	1998	Culture and IMS-PCR IS*900*	NA	90	1	1	Non-random sample. Only few herds included.	Grant et al., 2001
Mixed sheep and goats	Portugal /Lisbon	?	Parachek ELISA, CSL, Australia	?	66	18	27	Larger size flocks and one flock suspected of clinical disease selected.	Mendes et al., 2004
Mixed sheep and goats	Slovenia /all	2000-01	Pourquier ELISA		438	51	12	No distinction btw. sheep and goats	Ocepek et al., 2002
Sheep	Greece /Northern Sporades Islands	?	AGID & faecal culture	?	13	0	0	Information could not be verified	Minas et al., 1994
Sheep	Spain / Avila region	1996	AGID	> 18 months	38	11	29	Cross-sectional study of herds > 25 animals older than 18 months	Reviriego et al., 2000
Sheep	Switzerland /all	2002	Bulk tank milk IS*900* PCR	NA	63	15	24		Muehlherr et al., 2003
Sheep	United Kingdom /England, Wales, Northern Ireland	1998	Culture and bulk tank milk IS*900* IMS-PCR	NA	14	0	0	Non-random sample. Only few herds included.	Grant et al., 2001

Abbreviations: n = number of herds tested; T+ = number of herds test-positive; HAP = herd level apparent prevalence; NA = not applicable

3.3.1. Critical review of estimated true prevalences in sheep and goats on animal level

Among the relatively few prevalence studies in sheep and goats, all TPs were 0% (Table 7), although MAP was isolated from animals in most of the studies. The calculated TPs thereby appeared unlikely. TPs of 0% resulted from the combination of a low AP and an imperfect Sp. No study design and sampling frame was reported in the Austrian study (Khol et al., 2006), making appropriate inference on prevalences difficult. In Norway, only the 10 oldest animals (goats or sheep) in the herds were sampled (Mørk et al., 2003), thereby making the sample non-representative.

Mendes et al. (2004) reported that 1.7% of goats and sheep in larger flocks in the Lisbon region of Portugal were test-positive in ELISA, but the number of samples was not specified for each species, and species specific estimates were not given. Among 375 goats and 356 sheep in Croatia, only 2 sheep tested positive in an IDEXX ELISA (Cvetnic, 2002). It was not possible to determine whether these were false- or true-positive. The study comprising the highest number of animals was reported from Slovenia, where 12,578 sheep and goats were tested. The distribution between sheep and goats was not specified. A total of 440 (3.5%) tested positive, but because the test accuracy appear to be different for sheep and goats (Table 1), a TP could not be estimated.

To summarise, there were no studies that provided accurate and unbiased prevalence estimates among small ruminants. All APs were below 4%, but the TPs resulting from conversion using the Rogan-Gladen estimator were all 0%, irrespective that MAP had been isolated from goats and sheep in Norway and goats in Austria.

3.3.2. Critical review of estimated true prevalences in goats and sheep on herd level

The HAPs of MAP in small ruminants are summarised in Table 8. While all animal level results were deemed non-interpretable due to non-random sampling, aggregation of the results to herd level estimates would also be non-interpretable. Also, all TPs were 0% although several studies reported on knowledge that MAP infections did occur. From the Avila region in Spain, a HAP of 29% was reported from a random selection of sheep flocks, but insufficient information was available to calculate the HTP, because the number of animals sampled on each flock was not stated. However, Reviriego et al. (2000) stated that the AGID test used was 100% specific. Therefore, the HAP of 29% would result in a HTP of ≥29%.

Bulk tank milk sampling with subsequent direct IS*900* PCR have been used in several studies. Grant et al. (2001) conducted a study including only a few herds, which were randomly sampled. In Switzerland, the HAPs in goats and sheep were 23% and 29%, respectively. These estimates were probably underestimated, because not all infected animals shed MAP in milk, but there were no test accuracy estimates for the tests. Therefore, the guesstimate of MAP infected herds was ≥23%, but based on a sample of Swiss and Spanish small ruminants only. The remaining estimates were difficult to interpret.

3.4. Deer and llamas

Three studies have been carried out to assess the prevalence of MAP in farmed deer and llamas (Table 9). The accuracy of diagnostic tests was unknown for the target condition. The number of farms included in the studies was limited, and the farms from the Czech Republic (Machackova et al., 2004) did not constitute a random sample. It was unclear whether all fallow deer and red deer originated from the same farms. It was not possible to make inference on MAP prevalences in the general population of farmed deer and llamas, considering the low number of farms involved and thereby potential resulting effect of clustering within farms.

Table 9. Animal level apparent prevalences of *Mycobacterium avium* subspecies *paratuberculosis* infections in farmed deer and llamas in Europe based on information published in January 1990 to April 12, 2007

Species	Country /region	Period	Test	Age group	n	T+	AP (%)	Reference
Llama	Norway /all	2002	Faecal culture	All animals > 4 yrs	68	0	0	Mørk et al., 2003
Fallow deer	Czech Republic /all	1999-2001	Tissue culture	All	104	3	3	Machackova et al., 2004
Red deer	Czech Republic /all	1999-2001	Tissue culture	All	485	103	21	Machackova et al., 2004

4. DISCUSSION

Prevalence studies have been conducted across Europe, but very few studies provided comparable and interpretable estimates of herd and animal level prevalence of MAP infection in farmed animals. Due to the limitations of the studies, comparable within-herd prevalence estimates could only be established as guesstimates. The prevalence among cattle appeared to be approximately 20% and was at least 3–5% in several countries, whereas no animal level estimates were interpretable for sheep and goats. Between-herd prevalence estimates would also be guesstimates only, based on reports from only a few countries.

From these countries, the prevalence appeared to be >50% in cattle herds and >20% in sheep and goat flocks. In farmed deer and llamas, reasonable estimates could not be made. There were no studies with sufficient information to claim that a country or region was free of MAP infection.

Prevalence studies should enable decision makers to draw inferences with regard to control or surveillance of the infection, or for modellers to use as input parameters in epidemiological simulation models. A major critical issue in many studies reviewed in this report was the design of studies, where random sampling was not achieved, resulting in potentially biased and non-interpretable results. Considerations of study design should include definitions of target condition and target population. Both target population and condition should be defined by the decision makers. In most cases, the target population is all animals or herds of a given species in a given region. The study population must reflect the target population and should be obtained by a random sampling procedure. Selection of specific animal groups should be avoided if this selection results in a distribution of infected and un-infected animals, which is different from the distribution in the target population. Old non-infected animals are more likely to have survived because most animals have a progression of the infection in the age interval 2.2–4.4 years of age (Nielsen and Ersbøll, 2006). Sampling of old animals can thereby result in “healthy worker survivor” bias (Arrighi and Hertz-Picciotto, 1994), whereas sampling of young animals can result in low test accuracy, if the test cannot detect the infected animals. For MAP infections, ELISAs will usually not detect infected animals less than 2 years of age, as these have not developed antibodies to MAP (Nielsen and Ersbøll, 2006). Reduction of bias caused by a diagnostic test can be reduced by inclusion of age-specific test accuracy estimates (Sergeant et al., 2008).

Another requirement is that the diagnostic tests used should be able to discriminate between infected and non-infected animals. Current diagnostic tests are incapable of near-complete discrimination, emphasising the need for an appropriate test evaluation. Prior to a prevalence study, accuracy of the test to be used should be evaluated in the target population. Latent-class methodology (Nielsen et al., 2002; Wang et al., 2007) can be used as a tool for estimating the Se and Sp in situations, where infected and non-infected populations representative of the target population cannot be obtained. The target condition of MAP in the test-evaluation should be similar to the target condition in the prevalence study.

The apparent lack of valid prevalence studies in most parts of Europe would suggest that some decision makers may need new studies, if they wish to compare to situations in other countries. Such studies should be designed to aim at the target populations and the results should be interpreted with appropriate methods, ideally including test-evaluations of the tests used in the target populations with the case definition of paratuberculosis in mind. Most of the studies conducted so far could have been useful for across population comparisons had they included some form of test evaluation or validation. For the majority of the studies reported, more stringency should have been used in the study designs.

Should a prevalence study have been carried out appropriately, accurate reporting is also important. Many prevalence studies may not be reported in scientific literature, and have therefore not been included here. The minimum requirements in the reporting of a prevalence study are descriptions of: (a) target and study populations, including the age distribution; (b) the test used, including test accuracy, protocol, etc. (see Nielsen and Toft, 2008); (c) study period; (d) APs which are converted to TPs; and (e) separate estimates of

APs and TPs for each animal species. Novel approaches may be used for future prevalence studies. Test accuracy studies may not need to be performed separately, if latent class mixture models are used as demonstrated by Nielsen et al. (2007b). The advantage of that approach is that a test-evaluation is not needed, but disadvantages include that the models are not yet well-accepted in the veterinary society and thereby disbelief may predominate on the resulting estimates.

5. CONCLUSION

There are few valid MAP prevalence estimates available in Europe, partly due to problems with test accuracy in target populations, partly due to study designs unsuitable for prevalence inference. Therefore, a proper conclusion on the prevalence of MAP infections in Europe cannot be made. Instead, guesstimates were derived based on the TPs which appeared most likely to be valid. The true prevalence among cattle appeared to be approximately 20% and was at least 3–5% in several countries. Herd level prevalences were likely to be >50% in many countries. Some countries may have a lower prevalence, but the available information did not provide the necessary documentation. In sheep and goats, the within-flock prevalences could not be estimated. On flock level, the prevalences were thought to be >20% for both sheep and goats, based on studies from Switzerland and Spain. No reliable guesstimates could be derived for other farmed species.

6. ACKNOWLEDGEMENTS

This study was co-funded by the European Commission, 6th Framework Programme, project no. 023106 (FOOD) (ParaTBTools).

7. REFERENCES

Anon., 2000. Possible links between Crohn's disease and Paratuberculosis. Report of the Scientific Committee on Animal Health and Animal Welfare, European Commision, 76 pp.

Arrighi, H.M., Hertz-Picciotto, I., 1994. The evolving concept of the healthy worker survivor effect. Epidemiology 5, 189-196.

Atala, N., Akcay, E., 2001. Turkiye genelinde sgr paratuberkulozu prevalansnn ELISA ile arastrlmas (In Turkish, with English abstract). Etlik Veteriner Mikrobiyoloji Dergisi 12, 39-48.

Boelaert, F., Walravens, K., Biront, P., Vermeersch, J.P., Berkvens, D., Godfroid, J., 2000. Prevalence of paratuberculosis (Johne's disease) in the Belgian cattle population. Vet. Microbiol. 77, 269-281.

Bosshard, C., Stephan, R., Tasara, T., 2006. Application of an F57 sequence-based real-time PCR assay for *Mycobacterium paratuberculosis* detection in bulk tank raw milk and slaughtered healthy dairy cows. J. Food Prot. 69, 1662-1667.

Böttcher, J., Gangl, A., 2004. *Mycobacterium avium* ssp. *paratuberculosis*--combined serological testing and classification of individual animals and herds. J. Vet. Med. B. 51, 443-448.

Böttcher, M., 1997. Kulturelle und serologische Untersuchungen zur Pravalenz der Paratuberkulose im Regierungsbezirk Arnsberg (In German, with English abstract). Dissertation, Tierärztliche Hochschule Hannover, Germany, pp. 99.

Carter, M.A., 2007. An overview of the voluntary Bovine Johne's disease control program in the United States of America. Bulletin of International Dairy Federation, 410, 14-19.

Cetinkaya, B., Egan, K., Harbour, D.A., Morgan, K.L., 1996. An abattoir-based study of the prevalence of subclinical Johne's disease in adult cattle in south west England. Epidemiol. Infect. 116, 373-379.

Cetinkaya, B., Erdogan, H.M., Morgan, K.L., 1998. Prevalence, incidence and geographical distribution of Johne's disease in cattle in England and the Welsh borders. Vet. Rec. 143, 265-269.

Cetinkaya, B., Muz, A., Ertas, H.B., Öngör, H., Sezen, I.Y., Gülcü, H.B., 2000. Sut ineklerinde paratuberkuloz prevalansnn polimeraz zincir reaksiyonu (PZR) ile saptanmas (In Turkish, with English abstract). Turk Veterinerlik ve Hayvanclk Dergisi 24, 371-379.

Christensen, J., Gardner, I.A., 2000. Herd-level interpretation of test results for epidemiological studies of animal diseases. Prev. Vet. Med. 45, 830-106.

Corti, S., Stephan, R., 2002. Detection of *Mycobacterium avium* subspecies *paratuberculosis specific* IS*900* insertion sequences in bulk-tank milk samples obtained from different regions throughout Switzerland. BMC. Microbiol. 2,15.

Cvetnic, Z., 2002. Paratuberkuloza ovaca i koza u Republici Hrvatskoj (In Croatian, with English abstract). Praxis Veterinaria Zagreb 50, 255-260.

Donat, K., Eulenberger, K., Kampfer, P., 2005. Blutserologische Untersuchungen zur Verbreitung von *Mycobacterium avium* spp. *paratuberculosis* in sachsischen Rinderbestanden (In German, with English abstract). Tieraerztliche Umschau 60, 497-501.

Dreier, S., Khol, J.L, Stein, B., Fuchs, K., Gutler, S., Baumgartner, W., 2006. Serological, bacteriological and molecularbiological survey of paratuberculosis (Johne's disease) in Austrian cattle. J. Vet. Med. B 53, 477-481.

Franken, P., 2005. Paratuberculosis control in the Netherlands: the target and an overview of activities. *In*: Manning, E.J.B., Nielsen, S.S. (eds.), Proceedings of the 8th International Colloquium on Paratuberculosis, Copenhagen, Denmark, August 14-18, 2005, p. 15-19.

Gasteiner, J., Wenzl, H., Fuchs, K., Jark, U., Baumgartner, W., 1999. Serological cross-sectional study of paratuberculosis in cattle in Austria. J. Vet. Med. B 46, 457-466.

Grant, I.R., O'Riordan, L.M., Ball, H.J., Rowe, M.T., 2001. Incidence of *Mycobacterium paratuberculosis* in raw sheep and goats' milk in England, Wales and Northern Ireland. Vet. Microbiol. 79, 123-131.

Greiner, M., Gardner, I.A., 2000. Epidemiologic issues in the validation of veterinary diagnostic tests. Prev. Vet. Med. 45, 3-22.

Ikiz, S., Bagcigil, A.F., Seyyal, A.K., Ozgur, N.Y., Lgaz, A., 2005. Paratuberculosis in cattle in Turkey detected by PCR. Medycyna Weterynaryjna 61, 881-883.

Kennedy, D.J., 2007. Development in the approach to managing paratuberculosis in Australia. Bulletin of International Dairy Federation, 410, 8-13.

Kennedy, D.J., Benedictus, G., 2001. Control of *Mycobacterium avium* subsp. *paratuberculosis* infection in agricultural species. Rev. sci. tech. Off. Int. Epiz. 20, 151-179.

Khol, J.L., Stein, B., Dreier, S., Baumgartner, W., 2006. Paratuberculosis (Johne's disease) in small ruminants in Austria. Slo. Vet. Res. 43 (Supp 10), 129-130.

Kobayashi, S., Tsutsui, T., Yamamoto, T., Nishiguchi, A., 2007. Current surveillance and control program on paratuberculosis in Japan. Bulletin of International Dairy Federation 410, 30-33.

Kulkas, L., 2007. Plans for a paratuberculosis certification and eradication program in Finland. Bulletin of International Dairy Federation 410, 20-22.

Lillini, E., Bitoni, G., Gamberale, F., Cersini, A., 2005. Prevalence of bovine paratuberculosis in the Latium region (Italy). *In*: Manning, E.J.B., Nielsen, S.S. (eds.), Proceedings of the 8th International Colloquium on Paratuberculosis, Copenhagen, Denmark, August 14-18, 2005, p. 638-644.

Machackova, M., Svastova, P., Lamka, J., Parmova, I., Liska, V., Smolik, J., Fischer, O.A., Pavlik, I., 2004. Paratuberculosis in farmed and free-living wild ruminants in the Czech Republic (1999-2001). Vet. Microbiol. 101, 225-234.

Martin, S.W., Shoukri, M., Thorburn, M.A., 1992. Evaluating the health status of herds based on tests applied to individuals. Prev. Vet. Med. 14, 33-43.

Mendes, S., Boinas, F., Albuquerque, T., Fernandes, L., Afonso, A., Amado, A., 2004. Epidemiological studies on paratuberculosis in small ruminants in Portugal. Epidemiol. Sante Anim. 45, 61-71.

Meylan, M., Nicolet, J., Busato, A., Burnens, A., Martig, J., 1995. Etude de prevalence de la paratuberculose dans la region du Plateau de Diesse (In French, with English abstract). Schweiz. Arch. Tierheilkd. 137, 22-25.

Minas, A., Koutsoukou-Hartona, E., Papasavvas, M., Tsantas, H., 1994. Epidimiologiki erevna sta poimnia provaton kai aigon stis Voreies Sporades gia parafymatiosi kai proiousa pnevmonia (In Greek, with English abstract). Deltion tis Ellinikis Ktiniatrikis Etaireias 45, 25-30.

Muehlherr, J.E., Zweifel, C., Corti, S., Blanco, J.E., Stephan, R., 2003. Microbiological quality of raw goat's and ewe's bulk-tank milk in Switzerland. J. Dairy Sci. 86, 3849-3856.

Muskens, J., Barkema, H.W., Russchen, E., van Maanen, K., Schukken, Y.H., Bakker, B., 2000. Prevalence and regional distribution of paratuberculosis in dairy herds in The Netherlands. Vet. Microbiol. 77, 253-261.

Muskens, J., van Zijderveld, F., Eger, A., Bakker, D., 2002. Evaluation of the long-term immune-response in cattle after vaccination against paratuberculosis in two Dutch dairy herds. Vet. Microbiol. 86, 269-278.

Mørk, T., Heier, B.T., Alvseike, K.R., Lund, A., 2003. Overvåkings- og kontrollprogrammer for landdyr, fisk og skjell i Norge med vekt pa BSE-, salmonella- og paratuberkuloseprogrammene (In Norwegian, with English abstract). Norsk Veterinaertidsskrift 115, 707-717.

Nielsen, S.S., Ersbøll, A.K., 2006. Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows. J. Dairy Sci. 89, 4557-4566.

Nielsen, S.S., Toft, N., 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques. Vet. Microbiol. 129, 217-235.

Nielsen, S.S., Thamsborg, S.M., Houe, H., Bitsch, V., 2000. Bulk-tank milk ELISA antibodies for estimating the prevalence of paratuberculosis in Danish dairy herds. Prev. Vet. Med. 44, 1-7 (with Corrigendum in Prev. Vet. Med. 46, 297).

Nielsen, S.S., Grønbæk, C., Agger, J.F., Houe, H., 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53, 191-204.

Nielsen, S.S., Jepsen, Ø.R., Aagaard, K., 2007a. Control programme for paratuberculosis in Denmark. Bulletin of International Dairy Federation 410, 23-29.

Nielsen, S.S., Toft, N., Jørgensen, E., Bibby, B.M., 2007b. Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response. Prev. Vet. Med. 81, 290-305.

O'Doherty, A., O'Grady, D., O'Farrell, K., Smith, T., Egan, J., 2002. Survey of Johne's disease in imported animals in the Republic of Ireland. Vet. Rec. 150, 634-636.

O'Reilly, C.E., O'Connor, L., Anderson, W., Harvey, P., Grant, I.R., Donaghy, J., Rowe, M., O'Mahony, P., 2004. Surveillance of bulk raw and commercially pasteurized cows' milk from approved Irish liquid-milk pasteurization plants to determine the incidence of *Mycobacterium paratuberculosis*. Appl. Environ. Microbiol. 70, 5138-5144.

Ocepek, M., Posedi, J., Pislak, M., 1999. Prevalence of bovine paratuberculosis in Slovenia in 1997 and 1998. Zbornik Veterinarske Fakultete Univerza Ljubljana 36, 111-119.

Ocepek, M., Krt, B., Pate, M., Pogacnik, M., 2002. Seroprevalence of paratuberculosis in Slovenia between 1999 and 2001. Slo. Vet. Res. 39, 179-185.

Ott, S.L., Wells, S.J., Wagner, B.A., 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40, 179-192.

Petit, E., 2001. Enquete serologique sur la paratuberculose bovine menee dans l'Yonne lors de la campagne 98-99 (In French, with English abstract). Epidemiol. Sante Anim. 40, 23-39.

Reviriego, F.J., Moreno, M.A., Dominguez, L., 2000. Soil type as a putative risk factor of ovine and caprine paratuberculosis seropositivity in Spain. Prev. Vet. Med. 43, 43-51.

Robbi, C., Rossi, I., Nardelli, S., Rossi, E., Toson, M., Marangon, S., Vincenzi, G., Vicenzoni, G., 2002. Prevalenza di Paratubercolosi (Johnes disease) nella popolazione di bovine da latte della Regione Veneto (In Italian, with English abstract). Atti della Societa Italiana di Buiatria, 34, 283-288.

Rogan, W.J., Gladen, B., 1978. Estimating prevalence from the results of a screening test. Am. J. Epidemiol. 107, 71-76.

Sergeant, E.S.G., Nielsen, S.S., Toft, N., 2008. Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds. Prev. Vet. Med. 85, 92-106.

Sevilla, I., Aduriz, G., Garrido, J.M., Geijo, M.V., Juste, R.A., 2002. A preliminary survey on the prevalence of paratuberculosis in dairy cattle in Spain by bulk milk PCR. *In*: Juste, R.A., Geijo, M.V., Garrido, JM. (eds.), Proceedings of the 7th International Colloquium on Paratuberculosis, Bilbao, Spain, June 2002, p. 332-336.

Stärk, K.D., Frei-Staheli, C., Frei, P.P., Pfeiffer, D.U., Danuser, J., Audige, L., Nicolet, J., Strasser, M., Gottstein, B., Kihm, U., 1997. Häufigkeit und Kosten von Gesundheitsproblemen bei Schweizer Milchkühen und deren Kälbern (1993-1994) (In German, with English abstract). Schweiz. Arch. Tierheilkd. 139, 343-353.

Stephan, R., Buhler, K., Corti, S., 2002. Incidence of *Mycobacterium avium* subspecies *paratuberculosis* in bulk-tank milk samples from different regions in Switzerland. Vet. Rec. 150, 214-215.

Sternberg, S., Viske, D., 2003. Control strategies for paratuberculosis in Sweden. Acta Vet. Scand. 44, 247-249.

Sweeney, R.W., Whitlock, R.H., Buckley, C.L., Spencer, P., Rosenberger, A.E., Hutchinson, L.J., 1994. Diagnosis of paratuberculosis in dairy cattle using enzyme-linked immunosorbent assay for detection of antibodies against *Mycobacterium paratuberculosis* in milk. Am. J. Vet. Res. 55, 905-909.

Tharaldsen, J., Djonne, B., Fredriksen, B., Nyberg, O., Siguroardottir, O., 2003. The National Paratuberculosis Program in Norway. Acta Vet. Scand. 44, 243-246.

Vicenzoni, G., Filippi, L., Dolci, P., Corrò, M., Job, L., Robbi, C., Mutinelli, F., Marangon, S., 1999. Prevalenza del *Mycobacterium paratuberculosis* in bovine macellate in provincia di Verona (In Italian, with English abstract). Atti della Societa Italiana di Buiatria 31, 277-281.

Viske, D., Larsson, B., Engvall, A., Bölske, G., 1996. Paratuberculosis in Sweden. *In*: Chiodini, R.J., Hines III, M.E., Collins, M.T. (eds), Proceedings of the Fifth International Colloquium on Paratuberculosis, Madison, Wisconsin, USA, Sept. 29-Oct. 4, 1996, p. 143-146.

Wang, C., Turnbull, B.W., Gröhn, Y.T., Nielsen, S.S., 2007 Nonparametric estimation of ROC curves based on Bayesian models when the true disease state is unknown. Journal of Agricultural, Biological, and Environmental Statistics 12, 128-146.

Whitlock, R.H., Rosenberger, A.E., Sweeney, R.W., Spencer, P.A., 1996. *In*: Chiodini, R.J., Hines III, M.E., Collins, M.T. (eds), Proceedings of the Fifth International Colloquium on Paratuberculosis, Madison, Wisconsin, USA, Sept. 29-Oct. 4, 1996, p. 168-174.

Paper VIII

Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds

Sergeant ESG, Nielsen SS, Toft N

Published in
Preventive Veterinary Medicine 2008, 85, 92–106

Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds

E. S.G. Sergeant[1], S.S. Nielsen[2], N. Toft[2]

[1]International EpiLab, The National Veterinary Institute, Danish Technical University, Bülowsvej 27, 1790 Copenhagen V, Denmark
[2]Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark.

ABSTRACT

Paratuberculosis is a chronic infection affecting cattle and other ruminants. In the dairy industry, losses due to paratuberculosis can be substantial in infected herds and several countries have implemented national programmes based on herd-classification to manage the disease. The aim of this study was to develop a method to estimate the probability of low within-herd prevalence of paratuberculosis for Danish dairy herds. A stochastic simulation model was developed using the R® programming environment. Features of this model included: use of age-specific estimates of test-sensitivity and specificity; use of a distribution of observed values (rather than a fixed, low value) for design prevalence; and estimates of the probability of low prevalence (Pr_{Low}) based on a specific number of test-positive animals, rather than for a result less than or equal to a specified cut-point number of reactors.

Using this model, five herd-testing strategies were evaluated: (1) milk-ELISA on all lactating cows; (2) milk-ELISA on lactating cows ≤4 years old; (3) milk-ELISA on lactating cows >4 years old; (4) faecal culture on all lactating cows; and (5) milk-ELISA plus faecal culture in series on all lactating cows.

The five testing strategies were evaluated using observed milk-ELISA results from 19 Danish dairy herds as well as for simulated results from the same herds assuming that they were uninfected.

Whole-herd milk-ELISA was the preferred strategy, and considered the most cost-effective strategy of the five alternatives. The five strategies were all efficient in detecting infection, i.e. estimating a low Pr_{Low} in infected herds, however, Pr_{Low} estimates for milk-ELISA on age-cohorts were too low in simulated uninfected herds and the strategies involving faecal culture were too expensive to be of practical interest. For simulated uninfected herds, whole-herd milk-ELISA resulted in median Pr_{Low} values >0.9 for most herds, depending on herd size and age-structure. None of the strategies provided enough power to establish a high Pr_{Low} in smaller herds, or herds with a younger age-structure. Despite this, it appears as if the method is a useful approach for herd-classification for most herds in the Danish dairy industry.

INTRODUCTION

Paratuberculosis is an insidious, chronic infection primarily affecting cattle and other ruminants (Chiodini et al., 1984), caused by *Mycobacterium avium* subsp. *paratuberculosis* (Map). Herd-level prevalence varies from country to country in the range from very low (i.e. almost 0% in Finland and Sweden) to very high (e.g. >50% in Denmark and The Netherlands), as reviewed in Kennedy and Benedictus (2001). The production losses associated with Map-infection in herds where no measures to control infection are carried out

have been estimated at 300 €/cow-stall/year relative to herds without the infection (Kudahl, 2004). The cumulative losses can thus be quite significant if no control measures are in place and the prevalence is high.

Spread of Map among herds is primarily through purchase of infected animals (Sweeney, 1996; van Weering et al., 2005), so efficient testing strategies are required to identify low-risk animals or herds as a source of replacement animals (Kennedy and Benedictus, 2001). However, tests for paratuberculosis in individual animals are generally poor, especially lacking sensitivity, primarily due to the chronic nature of the infection and potential latency within herds and animals (Whittington and Sergeant, 2001). Sensitivity estimates for serum and milk-ELISAs for paratuberculosis vary depending on the age of the animal and stage of infection, but range from <10% in young animals to >80% in older animals with clinical disease or high levels of faecal shedding of Map (Dargatz et al., 2001; Jubb et al., 2004; McKenna et al., 2005; Nielsen and Toft, 2006). Specificity estimates also vary considerably, but are generally >95% and often >99% (Ridge et al., 1991; McKenna et al., 2005; Collins et al., 2005; Nielsen and Toft, 2006). As a result of the poor sensitivity, testing of individual cows from herds with unknown Map-infection status is unlikely to detect all infected animals and is thus of limited value for preventing spread of infection. However, herd-classification based on testing of whole herds or groups of animals in a herd, can be useful to reduce the risk of spread through purchased animals (Carpenter et al., 2004).

Various approaches to herd-classification have been used, but few have proven to be cost-effective. In the Netherlands, voluntary herd-classification based on annually repeated culture of pooled faecal samples was used in a programme established in 1998 (Benedictus et al., 2000). This approach was abandoned because farmers found it too costly, and herd-classification was unreliable, particularly considering the high cost. An alternative programme has been proposed with the objective of identifying herds that do not deliver more than 103 Map bacteria per litre in bulk tank milk (Franken, 2005).

In the USA, initial screening is based on a serum ELISA on a minimum of 30 animals, with subsequent confirmation using culture of faecal samples for the classification of herds. Progression to higher levels in the programme requires larger sample sizes, to provide 95% confidence of detecting a prevalence ≥2%.The programme aims at identifying herds with a low-risk of Map-infection (Bulaga, 1999). A similar programme operates in Australia, with up to 300 animals screened per herd (depending on herd size) by serum ELISA, with follow-up faecal culture on ELISA positives (Kennedy et al., 2002). These approaches provide increasing confidence in herd-freedom from paratuberculosis with each additional herd-test, but may fail to detect herds with a low prevalence of infection, particularly in the US programme, where entry into the programme can be based on testing only 30 randomly selected cows ≥2nd parity (Wells et al., 2002).

The test-strategies used in these programmes do not fully account for changes in test-sensitivity with age (Nielsen and Toft, 2006). Instead, only animals >3 years of age are tested in the US programme and >2 years for the initial test and >4 years in subsequent tests in the Australian programme, to take advantage of higher test-sensitivity in older animals. These programmes also rely on follow-up testing with faecal culture to confirm infection. However, although faecal culture is highly specific, it is less sensitive, but time-consuming and costly, thus adding further to testing costs.

Simulation studies have also been used in an attempt to identify cost-effective testing strategies for detection of Map-infected herds (Weber et al., 2004; Tavornpanich et al., 2006;

Tavornpanich et al., 2008). These studies showed that targeted sampling of smaller numbers of older animals and using a cheaper test such as serology was the most cost-effective of the individual animal sampling strategies for detecting infection at a specified design prevalence. Culture of environmental samples was more cost-effective than individual animal sampling strategies, but provides no information on which animals are likely to be excreting Map organisms (Tavornpanich et al., 2008). However, environmental sampling and targeted sampling of small numbers of animals provide only limited information to assist herd management, compared to sampling of larger numbers of individual animals.

These studies, as well as the US and Australian programmes are based on testing to provide a specified level of confidence of detecting paratuberculosis, if it were present at a specified within-herd prevalence (herd-sensitivity), rather than on the probability that true prevalence is less than the specified design prevalence. They also assume 100% specificity for the testing regimen used, rather than allowing and adjusting for false-positive reactions (Martin et al., 1992; Cameron and Baldock, 1998). A herd-classification scheme that is based directly on the probability of low prevalence of infection of the herd and which allows for imperfect specificity and false-positive reactions would therefore provide a useful alternative to existing approaches.

The objectives of our study were to:

a. develop a method for estimation of the probability of low prevalence of infection with *Mycobacterium avium* subsp. *paratuberculosis* in Danish dairy herds based on individual animal-testing results using imperfect tests; and
b. compare the cost and effectiveness of alternative testing strategies for estimating the probability of low prevalence.

MATERIALS AND METHODS

Model overview

"ParaFree", a stochastic simulation model to estimate the probability of a low (design) within-herd prevalence for paratuberculosis, was developed using the R® programming environment (v 2.3.1, 2006). Novel features of this model include the analysis of real (or simulated) animal-testing data, use of age-specific estimates of test-sensitivity and specificity, and estimates of the probability that true within-herd prevalence is less than a specified design prevalence (Pr_{Low}). ParaFree estimates are also based on a specific (observed) number of test-positive animals, rather than on a hypothetical result less than or equal to a specified cut-point number of positives.

Inputs to the ParaFree model are probability distributions of age-specific sensitivity and specificity of individual milk-ELISA and faecal-culture tests for paratuberculosis in tested cattle and the assumed design prevalence, at which Pr_{Low} is calculated. Outputs of the model are corresponding distributions for Pr_{Low} and estimated true prevalence in the herd (see Fig. 1).

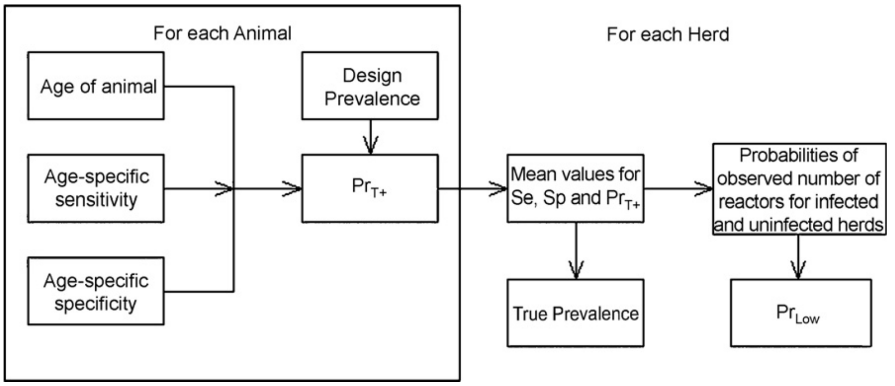


Fig. 1. Procedure for calculating probability of low prevalence (Pr_{Low}) based on age-specific sensitivity and specificity for each animal and design prevalence values.

Design Prevalence

An essential requirement for the estimation of Pr_{Low} is an estimate of the minimum prevalence of disease within infected herds for classification as low prevalence, commonly termed the threshold or design prevalence (DP) (Cameron and Baldock, 1998; Martin et al., 2007). Pr_{Low} is therefore the probability that the true prevalence (TP) of infection is less than the specified design prevalence.

DP is commonly used for studies to demonstrate freedom from disease. It is usually determined a priori and is often a fixed low prevalence reflecting the biology of the disease and the desire to detect infection at an early stage (Cameron and Baldock, 1998; Martin et al., 2007). Because the true prevalence of paratuberculosis in infected Danish dairy herds is generally high (mean 34%, range 4–85%) (Nielsen et al., 2007), we chose to use a distribution of observed values for design prevalence, rather than a fixed, low value. For each herd and iteration, an estimate of prevalence was drawn at random from a distribution of 100 estimates of true prevalence for the latest period of testing (Sampling period 3) described by Nielsen et al. (2007).

Sensitivity and specificity of ELISA and faecal culture

For each animal tested, estimates of sensitivity (Se) and specificity (Sp) were calculated from age-specific probability distributions based on data from Nielsen and Toft (2006) and summarised in Table 1. In that study, the probability of a test-positive response was estimated as a function of age and infection status using non-parametric regression for milk-ELISA and FC. Mean and 95% confidence intervals were calculated for age as a continuous variable ranging from 2 to 5 years. Mean milk-ELISA Se ranged from 0.06 in 2-year-old animals to 0.5 in 5-year-olds, milk-ELISA Sp ranged from 0.997 (2 years) to 0.93 (5 years) and faecal-culture Se ranged from 0.05 (2 years) to 0.2 (5 years). Sp for faecal culture was assumed to be 1 for all ages. Parameters for a Beta probability distribution can be calculated using estimates of the mode and a percentile of the required distribution (Suess et al., 2002). For each age value in the data, parameters for corresponding Beta distributions for Se and Sp were estimated based on the mean (assumed to approximate the mode for this estimation) and 2.5% or 97.5% percentile. The mean was used to estimate the mode because modal values were not available from the analysis of Nielsen and Toft (2006) and

visual inspection of the data suggested that confidence intervals were reasonably symmetrical. The 2.5% percentile was used for ELISA Sp estimates and the 97.5% percentile for ELISA and FC Se estimates because these values gave slightly wider distributions as the parameters approached 0 or 1. Selected examples of Se and Sp values and their corresponding distribution parameters are shown in Table 1.

Table 1. Mean and 95% confidence limits for age-specific milk-ELISA sensitivity and specificity and FC sensitivity estimates for selected ages of Danish dairy cattle (from Nielsen and Toft, 2006) and parameters for the corresponding Beta probability distributions

Age	Mean	95% CI	Alpha parameter	Beta parameter
ELISA Se				
2	0.058	0.039 – 0.085	24.41	381.21
2.5	0.201	0.183 – 0.221	340.27	1349.64
3	0.295	0.275 – 0.315	592.37	1414.27
3.5	0.336	0.316 – 0.357	661.18	1305.64
4	0.395	0.372 – 0.419	650.23	995.39
4.5	0.434	0.407 – 0.461	584.12	761.47
5	0.496	0.439 – 0.554	141.16	143.42
ELISA Sp				
2	0.997	0.991 – 0.999	966.83	3.91
2.5	0.992	0.989 – 0.995	4219.45	35.02
3	0.974	0.969 – 0.978	5369.37	144.3
3.5	0.947	0.94 – 0.953	4430.59	248.91
4	0.943	0.936 – 0.95	4899.1	297.07
4.5	0.924	0.914 – 0.932	2710.22	223.84
5	0.926	0.907 – 0.942	808.03	65.49
FC Se				
2	0.046	0.03 – 0.072	18.93	372.85
2.5	0.105	0.092 – 0.12	193.42	1641.15
3	0.157	0.143 – 0.172	364.89	1954.88
3.5	0.154	0.14 – 0.168	437.07	2396.55
4	0.148	0.134 – 0.164	322.93	1854.27
4.5	0.171	0.154 – 0.189	317.69	1536.3
5	0.213	0.173 – 0.258	78.79	288.42

For each animal sampled, age at sampling was calculated in decimal years and random values for Se and Sp for milk-ELISA and Se for FC were then generated from the appropriate age-specific Beta distribution for that animal. Where an animal's age was not represented in the Se or Sp data, the values for the next lowest age were used. For animals less than 2 years old (the minimum age in the Se/Sp dataset), Se and Sp were assumed to be 0 and 1, respectively. FC specificity was assumed to be 1, because the test was being used for herd-level rather than animal classification and because all positive cultures were confirmed by IS*900* PCR, which was assumed to be effectively 100% specific for Map when used as a herd-test. For testing with milk-ELISA and faecal culture in series, combined estimates of Se and Sp were calculated, assuming that the two tests were independent.

Calculating Pr_{Low} and true prevalence

The individual estimates of Se and Sp were used to calculate the probability of a positive test result (PrT+) for each animal, assuming the herd was infected at the design prevalence of DP: $Pr_{T+}=Se\times DP+(1-Sp)\times(1-DP)$

Mean values for Se, Sp and PrT+ were then used to calculate the probabilities of obtaining the specified number of test-positives for each herd, firstly assuming that it was infected at the design prevalence [P(x = r|inf)] and secondly assuming that it was uninfected [P(x = r|uninf)], using the binomial probability density function:

$$P(x = r \mid inf) = \binom{n}{x} Pr_{T+}^{x} (1 - Pr_{T+})^{(n-x)}$$

and

$$P(x = r \mid uninf) = \binom{n}{x} (1 - Sp)^{x} Sp^{(n-x)}$$

where *n* is the number tested, *r* the number of reactors.

Given estimates of herd-sensitivity and herd-specificity, the probability of freedom for each herd is the same as the negative predictive value of the herd-test (Martin et al., 2007). The above probabilities are analogous to 1—herd-sensitivity [P(x = r|inf)] and herd-specificity [P(x = r|uninf)] and can be substituted into the formula for negative predictive value to calculate Pr_{Low}:

$$Pr_{Low} = \frac{P(x = r \mid uninf) \times (1 - Pr_{Inf})}{(P(x = r \mid uninf) \times (1 - Pr_{Inf}) + Pr_{Inf} \times P(x = r \mid inf)}$$

where PrInf is the prior probability that the herd is infected. Based on previous testing (unpublished), the herd-prevalence of paratuberculosis in the Danish dairy industry is thought to be quite high (>50%). Therefore, a Beta distribution was used for PrInf, assuming a 5th percentile of 0.5 and a most likely value of 0.75, i.e. Beta (9.64, 3.88).

True prevalence of infection for each herd and iteration was estimated based on the apparent prevalence (AP) and mean values of Se and Sp for the herd (Rogan and Gladen, 1978), limited to between 0 and 1.

$TP = (AP + SP - 1)/(Se + Sp - 1)$

Testing strategies

ParaFree was used to estimate Pr_{Low} and TP for selected testing strategies in 19 herds for which testing data were available. All simulations were run for 1000 iterations to generate probability distributions for model outputs. Strategies modelled in this analysis were: (1) individual milk-ELISA on all lactating cows; (2) individual milk-ELISA on lactating cows less than or equal to 4 years old; (3) individual milk-ELISA on lactating cows greater than 4 years old; (4) individual faecal culture (FC) on all lactating cows; and (5) milk-ELISA plus FC in series on all lactating cows.

For Strategies 1–3, Pr_{Low} estimates were calculated for the observed test results for the 19 herds as well as for simulated test results assuming that the herds were not infected (see below). For Strategies 4 and 5, the tests were assumed to be 100% specific at the herd-level, so that for these tests, a single positive test result would result in a Pr_{Low} of 0. Therefore, analyses for these strategies were only undertaken for simulated uninfected herds, to explore the effects of herd size and age-distribution and to allow comparison of Pr_{Low} values with the other strategies.

Simulated uninfected herds

Pr_{Low} was estimated for the 19 herds using simulated data assuming that the herds were uninfected, to evaluate the methodology in genuinely uninfected herds. For this analysis, age-specific Sp was estimated for each animal as described above and compared to a random value generated between 0 and 1. If the random value exceeded the estimated Sp the animal was assumed to be test-positive. The total number of simulated test-positives for the herd was used as the assumed test result for the calculation of Pr_{Low} assuming the herd was uninfected.

Herd testing data

Test results for 19 Danish dairy herds from an intervention project against paratuberculosis and *Salmonella* Dublin were analysed to illustrate and evaluate the methodology. These herds were not a random selection of herds, but serve as examples for the model. Some herds considered paratuberculosis a problem and some considered *Salmonella* Dublin a problem. All lactating cows in each herd were tested four times per year with a milk antibody ELISA and once per year all cows in the herd were tested by cultivation of faecal samples on Herrolds egg yolk medium (HEYM). The tests were carried out in the period August 2003–October 2006. Only the first whole-herd milk-ELISA test for each herd was included in the analysis to minimise bias due to previous testing activity.

To allow calculation of age-specific sensitivity and specificity estimates, the age of each animal was calculated in decimal years as the test date minus the birth date, divided by 365. Animals for which the birth date was not available were excluded from the calculations.

Milk samples were collected from all lactating cows in each of the herds for routine testing in the Danish milk-recording scheme and milk-ELISA tests were undertaken at Steins Laboratory (Holstebro, Denmark) using a previously described method (Nielsen, 2002). Faecal samples were collected from all cows in the herd and cultured on HEYM at Steins Laboratory as described in Nielsen et al. (2004).

Table 2. Unit costs for milk-ELISA and faecal culture (FC) testing used to estimate testing costs for each testing strategy in 19 Danish dairy herds

Sampling costs	ELISA	FC	FC with Vet
Mileage rate	0	€ 0.45	€ 0.45
Mean km/herd	0	200	200
FC(Travel)	0	€ 632	€ 85
FC Samples per man hour	0	15	15
Cost per man hour	0	€ 50	€ 175
Sample bottles/gloves (per sample)	0	€ 0.60	€ 0.60
Freight (per herd)	0	€ 33	€ 33
Lab cost per sample	€ 3	€ 21	€ 21
Lab handling for sub-sampling per herd (Strategies 2 and 3 only)	€ 14		

[a] Faecal culture of ELISA positive animals (where required) was assumed to be undertaken by a veterinarian at an increased cost, compared to routine faecal culture undertaken by technicians.

Testing costs

Testing costs for each herd and strategy were estimated based on unit costs in Table 2. There was no cost for sample collection for milk-ELISA testing because this was undertaken on routinely collected samples. Strategies 2 and 3 included a € 14 handling cost for each herd, for the laboratory to separate the relevant samples for testing under these strategies. For the FC strategy (Strategy 4), technicians were assumed to collect faecal samples from all lactating cows in the herd, while for the milk-ELISA plus FC strategy (Strategy 4), a veterinarian was assumed to collect faecal samples from milk-ELISA positives, at a higher cost.

RESULTS

Descriptive summary for 19 herds

Numbers of animals tested and the number of milk-ELISA-positive and faecal-culture test results for the 19 herds are summarised in Table 3. The mean number of animals tested per herd was 110 (range 57–204). In total, 16% of the tested animals were milk-ELISA-positive (range between herds 1.6% and 39%). The mean age of animals was 4.2 years, ranging from 3.6 to 5.2 years between herds. One herd (Herd 3) has not had Map isolated by faecal culture, despite the relatively high percentage of test-positive animals. The remaining herds have all had Map isolated on one or more occasions.

Table 3. Summary of testing results and mean age of cows tested for 19 Danish dairy herds

Herd	Year tested	No. Tested	No. Positive	% Positive	Mean Age	Years FC positive
1	2003	189	74	39	4.0	2003, 2004, 2005, 2006
2	2004	177	26	15	4.3	2006 [a]
3	2004	112	15	13	4.1	– [a]
4	2003	80	15	19	3.9	2003, 2004, 2005, 2006
5	2003	112	12	11	4.3	2003, 2004, 2005, 2006
6	2003	98	18	18	4.3	2003, 2004, 2005, 2006
7	2003	68	16	24	3.6	2003, 2004, 2005, 2006
8	2003	94	15	16	4.2	2003, 2004, 2005, 2006
9	2004	65	2	3	3.8	2004, 2005
10	2003	86	5	6	4.1	2003, 2005, 2006
11	2003	128	23	18	4.3	2003, 2004 [b]
12	2003	108	6	6	3.9	2003, 2004, 2005, 2006
13	2004	204	10	5	4.3	2003, 2005, 2006
14	2003	71	17	24	5.2	2003, 2004, 2005, 2006
15	2004	111	34	31	4.2	2003, 2004, 2005, 2006
16	2003	95	7	7	3.8	2003, 2004, 2005, 2006
17	2004	57	16	28	4.0	2003, 2004, 2005, 2006
18	2004	107	24	22	4.7	2003, 2004, 2005
19	2005	123	2	2	3.7	2006 [c]
Total		2085	337	16	4.2	

[a] Whole-herd faecal culture in 2004 and 2005 all negative

[b] Ceased production in 2004

[c] Only tested in 2005 and 2006

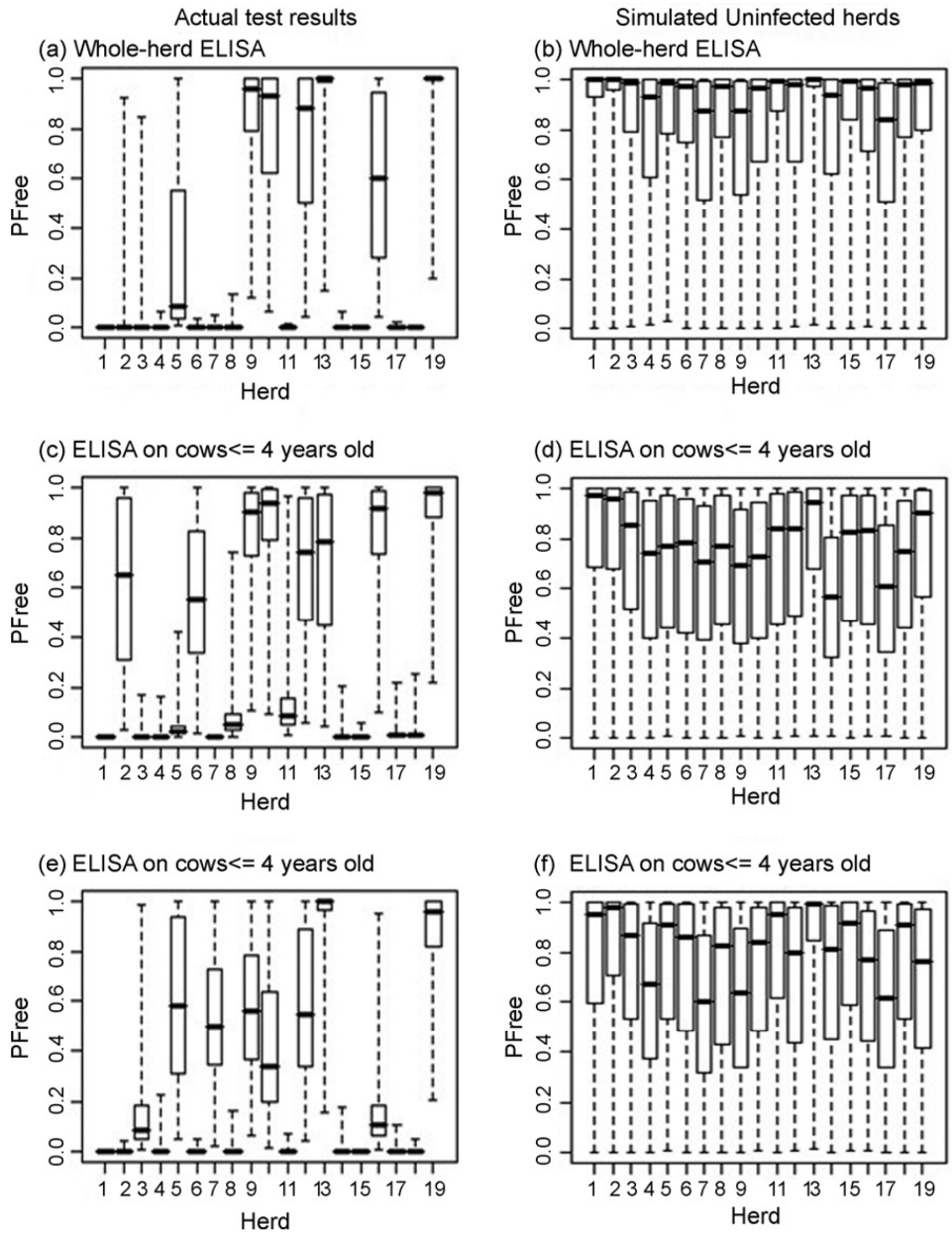


Fig. 2. Boxplots of the distribution of values for the probability of low prevalence (Pr_{Low}) for milk-ELISA testing in 19 Danish dairy herds. Solid bar is the median value, box represents 25th–75th percentiles and whiskers show range of values. Left column is for actual testing results (a, c and e), right column (b, d and f) for simulated uninfected herds.

Probability of low prevalence

Pr_{Low} estimates for Strategies 1–3 are summarised in Fig. 2, for the actual test results and simulated uninfected herds. When the actual test results for whole-herd milk-ELISA were

analysed, only two herds had a high probability of freedom (Herds 13 and 19, median >0.999 and 25th percentile >0.98), while 12 herds had Pr_{Low} close to zero and the remaining five herds had median values ranging from 0.09 to 0.95. Conversely, for simulated uninfected herds, whole-herd milk-ELISA resulted in median Pr_{Low} values >0.9 for most herds and 25th percentile values ranging from 0.5 to 0.95, depending on herd size and age-structure. Pr_{Low} values were generally lower in smaller herds (≤80 cows), compared to larger herds.

Actual testing results for young animals produced considerably lower Pr_{Low} values, with seven herds having median values for Pr_{Low} less than 0.01 and only three with median values greater than 0.9 (0.94, 0.92 and 0.98 for Herds 10, 16 and 19, respectively). Testing of only older animals produced somewhat higher estimates than those for younger animals but still lower than for whole-herd-testing. For actual testing results, ten herds had median values less than 0.01 and two herds had median values greater than 0.9 (0.999 and 0.96 for herds 13 and 19, respectively). A similar decrease was also apparent for testing of only younger or older animals for the simulated uninfected herds.

Median Pr_{Low} values for FC, assuming zero positive results, were generally similar to or slightly higher than those for the whole-herd milk-ELISA in simulated uninfected herds, while milk-ELISA and FC in series produced considerably lower values (Fig. 3), due to the substantially lower animal-level sensitivity of the serial testing regimen.

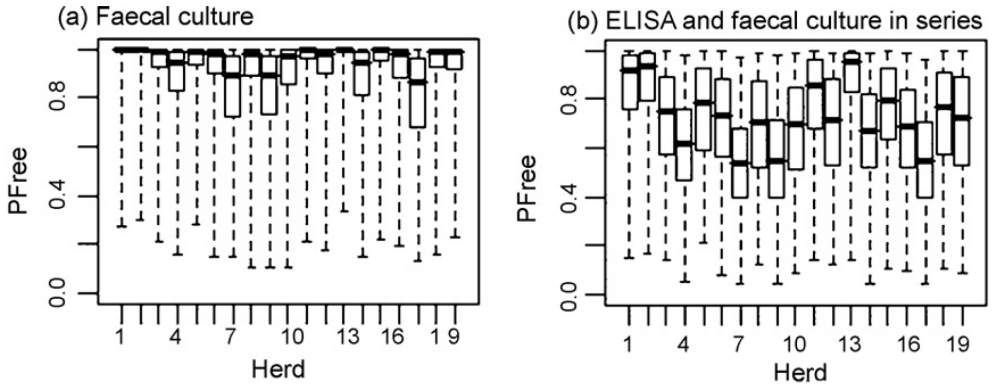


Fig. 3. Boxplots of the distribution of values for the low prevalence (Pr_{Low}) for faecal culture testing and faecal culture plus milk-ELISA in series in 19 simulated uninfected Danish dairy herds. Solid bar is the median value, box represents 25th–75th percentiles and whiskers show range of values.

True Prevalence

It was only possible to estimate TP for whole-herd milk-ELISA testing. True prevalence was not estimated for testing of younger or older animals because such estimates would not be representative of the herd. For FC and milk-ELISA plus FC, a single positive result was sufficient to classify the herd as infected, and no attempt was made to model the numbers of positives that might occur. For the whole-herd milk-ELISA and actual testing results, median TP ranged from 0 (3 herds) to 1 (1 herd), with most between 0.05 and 0.85, similar to the distribution used for DP (Fig. 4).

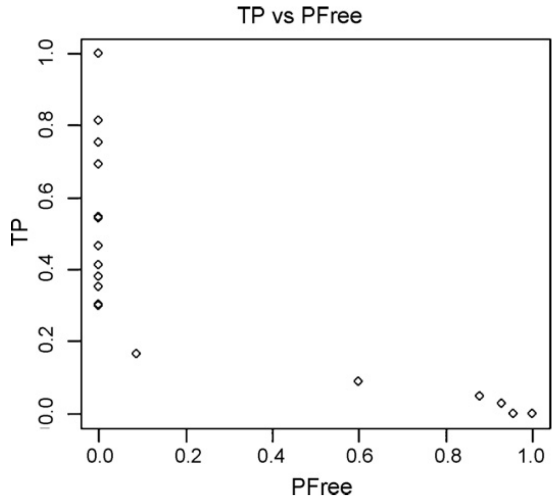


Fig. 4. Plot of median value for estimated true prevalence (TP) against median value for estimated low prevalence (Pr_{Low}) in 19 Danish dairy herds based on actual testing results.

Theoretically, herds with a high probability of low prevalence should have a low true prevalence and vice versa. To investigate this, a comparison of median values for TP and Pr_{Low} is shown in Fig. 4. Herds with a high TP (median > 0.2) had Pr_{Low} = 0, while four herds with a high Pr_{Low} (median > 0.9) had median TP values less than 0.03, including 3 with TP = 0.

Testing costs

Testing costs for the five testing strategies are summarised in Table 4. On average, FC was more than eight times the cost and milk-ELISA plus FC was about three times the cost of whole-herd milk-ELISA alone. The higher costs were due to a combination of sample collection costs, which were not required for milk-ELISA, and higher laboratory costs for FC compared to the ELISA. Milk-ELISA on younger animals was slightly more than half the cost of whole-herd milk-ELISA and slightly more than milk-ELISA on older animals.

Table 4. Summary of estimated testing costs for five testing strategies in 19 Danish dairy herds

	Whole-herd milk-ELISA	Milk-ELISA $A \leq 4$ years	Milk-ELISA $A > 4$ years	Faecal Culture	Milk-ELISA plus Faecal Culture
Mean cost per herd (€)	329	183	146	2859	1042
Minimum cost (€)	171	101	83	1544	385
Maximum cost (€)	612	365	323	5209	3152

DISCUSSION

In this study we have developed a model for estimating the probability of low prevalence of infection for individual herds using imperfect tests, without follow-up confirmatory testing of positive animals to classify the herd. It is the first study of its type that the authors are aware of that makes use of age-specific test-sensitivity and specificity to estimate the probability of

low infection prevalence. The method could also be easily adapted for demonstration of infection "freedom" by selecting a suitably low design prevalence.

Analyses of the actual test results from the 19 herds generally gave low estimates of Pr_{Low}, with only two of the 19 herds analysed having Pr_{Low} estimates close to 100%. These two herds have been shown to be infected, based on positive faecal-culture results. One of the herds had done some previous testing and might have biased the results by culling of reactors, while the other herd was faecal-culture negative at the time of testing in 2005 and has only been confirmed as infected during 2006. TP estimates in these herds were low, consistent with a high Pr_{Low} for the DP distribution used. One herd from which Map has not been isolated had quite low Pr_{Low} values and high TP, suggesting that some other factors have caused an unusually high false-positive milk-ELISA rate in this herd. These conflicting results suggest that while this approach appears to be feasible, actual results must still be treated with caution and interpreted in the light of other information available on individual herds.

An important component of a classification program is to identify low-risk or preferably disease free herds. The absence of good, cheap, fast and sensitive individual tests and the high prevalence of paratuberculosis at the herd-level in Denmark prohibited the identification and use of a large set of truly paratuberculosis free herds. Thus simulated data from the 19 herds, assuming they were truly uninfected, were used. This analysis showed that smaller herds or herds with a majority of younger animals, where milk-ELISA sensitivity is lower, are likely to have a low Pr_{Low}, despite being truly uninfected. However, most herds had a median $Pr_{Low} > 0.9$. It appears by visual assessment of Fig. 2 and Fig. 3 that whole-herd faecal culture performed better than whole-herd milk-ELISA, which performed better than milk-ELISA plus faecal culture in giving high Pr_{Low} estimates from truly uninfected herds. Considering that the cost of whole-herd faecal culture is about eight times the cost of whole-herd milk-ELISA, and there is a 12 week time delay in test result of faecal culture compared to milk-ELISA, the whole-herd milk-ELISA option must be preferred.

All of the five testing strategies lacked sufficient power to establish a high Pr_{Low} in small herds, but from our point of view, whole-herd milk-ELISA is to be preferred over the alternatives; the added cost compared to the age-cohort milk-ELISA strategies are justified when considering the improved power, i.e. consistently higher Pr_{Low}; the cost of whole-herd faecal culture makes it too expensive to be considered a realistic option; and follow-up faecal culture does not provide additional information, making it clearly inferior. Considering the range of Pr_{Low} estimates from both simulated uninfected herds and from actual milk-ELISA testing results suggests that, at least for the current population of Danish dairy herds, whole-herd milk-ELISA (or any of the other strategies) is better for ruling in paratuberculosis than ruling out paratuberculosis.

The use of age-specific sensitivity and specificity estimates for the milk-ELISA and FC in this study was important, as it allowed us to account for age-structure of herds when estimating Pr_{Low}. For example, estimated milk-ELISA Se increases from 0.06 in cows at 2 years of age to 0.2 at 2.5 years and 0.3, 0.4 and 0.5 at 3, 4 and 5 years, respectively. Therefore, an uninfected herd with a high proportion of older animals is likely to have a higher average sensitivity and hence higher Pr_{Low}. This was apparent in the simulated uninfected herds, where Herd 17 (mean age = 4.0) had a similar median Pr_{Low} to Herds 7 and 9 (mean ages = 3.6 and 3.8, respectively) (Fig. 2) despite having fewer animals tested (57 compared to 68 and 65 for Herds 7 and 9, respectively).

The age-specific specificity estimates used in this analysis were estimated in infected herds, and hence could under-estimate the true values (Nielsen and Toft, 2006). Thus, the Pr_{Low} estimates based on these would also be underestimated. Use of paratuberculosis-free herds for the test evaluation would have prevented this potential bias. However, there are no published estimates from free herds and the specificity estimates in the above study were made in animals with eight negative faecal-culture tests over a 2-year period, so that any bias in these estimates is assumed to be small. Similarly, use of sensitivity estimates from low-prevalence herds would have provided valuable additional information but such estimates were not available.

Sensitivity and specificity of the milk-ELISA and sensitivity of FC were determined in the Danish dairy cattle population, using the same tests as considered in this study. Therefore the results presented here are specific to this population and the values used are appropriate. In addition, Beta probability distributions were used to represent these values, allowing for uncertainty about the estimates. If suitable estimates of test performance were available the method could be easily adapted to other tests and populations.

Pr_{Low}, rather than herd-sensitivity, was chosen as the main output for ParaFree for several reasons. Firstly, it is an easier concept to explain and for producers to understand. Secondly, this methodology allows calculation of Pr_{Low} for an observed test result, rather than for less than or equal to a hypothetical cut-point number of positives, as is the case for herd-sensitivity. Finally, as repeated testing is undertaken (for example annually), Pr_{Low} can be updated over time by using the posterior estimate of Pr_{Low} from 1 year to generate an appropriate probability distribution for the prior probability of infection for the next year, rather than using a distribution based on industry estimates. Some adjustment would be necessary to reflect the decay of information in historical data and lack of independence between tests (Martin et al., 2007), however, this procedure might be one means by which small herds, where the Pr_{Low} estimated by a single test round inevitably will be low, could improve Pr_{Low} over time.

An important aspect of this study is that we have used a distribution of previously observed values for the design prevalence, rather than a fixed, low value. A distribution of observed values was chosen because this reflects the known distribution of within-herd prevalence of paratuberculosis in Danish dairy herds. The appropriateness of this distribution is also supported by the TP values estimated for the 19 herds in this study. In addition, considering the poor sensitivity of milk-ELISA and faecal culture, we believe that use of a low-design prevalence (for example 2% as used in Australia and the United States) would result in an inability for the method to differentiate between uninfected or low-prevalence herds and high-prevalence herds and would be of no practical use.

Use of realistic design prevalence values also gives rise to a potential source of bias in Pr_{Low} estimates. Because paratuberculosis is known to be endemic in the Danish dairy industry, a small proportion of herds are likely to be infected at levels below the design prevalence. However, these herds have not been taken into account in calculating $P(x = r|uninf)$, which is based on specificity of the test in truly uninfected herds only. This approach was used because of the difficulty in reliably estimating the proportion of herds in this category, the distribution of prevalence in these herds and the probability of the specified test results occurring in such herds. In addition, we believe that any bias introduced is likely to be of little consequence in the interpretation of results because: (1) the proportion of herds in this category is thought to be small; and (2) quite large changes in $P(x = r|uninf)$ result in only

small changes in Pr_{Low} as Pr_{Low} approaches 1, where potential bias would be of most concern.

This study intentionally focused on individual animal tests using the milk-ELISA and faecal culture. Herd-based tests such as PCR or culture of bulk milk or environmental samples could also be used as the basis for herd-classification at a considerably lower cost. These tests were not considered because: (1) they have not been evaluated under Danish conditions; (2) correlation of results with prevalence in infected herds is still not well characterised; and (3) they provide no information on individual animals to assist in decision-making and management of paratuberculosis by the herd manager.

CONCLUSION

This study has demonstrated that, despite the poor diagnostic properties of diagnostic tests for paratuberculosis, it is possible to derive potentially useful estimates of probability of herds having a low prevalence of paratuberculosis for the Danish dairy industry. Among the evaluated strategies whole-herd milk-ELISA on all lactating cows was the most cost-effective. None of the strategies had enough power to clearly establish freedom of infection in smaller herds or herds with a younger age-structure, however, repeated testing and additional information (such as movement patterns) can potentially be used for updating Pr_{Low} estimates, and thus provide a means for smaller herds to be classified.

Implementation of the methodology presented in this study in a dairy industry, such as the Danish, could best be achieved by setting classification criteria for herds based on median or percentiles of Pr_{Low} (for example 25th percentile >0.9), perhaps with multiple levels and provision for periodic updating of classification based on subsequent tests.

ACKNOWLEDGEMENTS

E.S.G. Sergeant conducted this study as Guest Scientist at the International EpiLab. The project was supported by the Danish Directorate of Food, Fisheries and Agro Business (DFFE, J.nr. 3412-05-1092).

REFERENCES

Benedictus, G., Verhoeff, J., Schukken, Y.H., Hesselink, J.W., 2000. Dutch paratuberculosis programme history, principles and development. Vet. Microbiol. 77, 399-413.

Bulaga, L.L., 1999. US Voluntary Johne's Disease Herd Status Program for Cattle. In: Manning EJB, Collins MT (Eds.), Proceedings of the 6th International Colloquium on Paratuberculosis. Madison, Wisconsin. pp 39-47.

Cameron, A.R., Baldock, F.C., 1998. Two-stage sampling in surveys to substantiate freedom from disease. Prev. Vet. Med. 34, 19-30.

Carpenter, T.E., Gardner, I.A., Collins, M.T., Whitlock, R.H., 2004. Effects of prevalence and testing by enzyme-linked immunosorbent assay and fecal culture on the risk of introduction of *Mycobacterium avium* subsp. *paratuberculosis*-infected cows into dairy herds. J. Vet. Diagn. Invest. 16, 31-38.

Chiodini, R.J., Van Kruiningen, H.J., Merkal, R.S., 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74, 218-262.

Collins, M.T., Wells, S.J., Petrini, K.R., Collins, J.E., Schultz, R.D., Whitlock, R.H., 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12, 685-692.

Dargatz, D.A., Byrum, B.A., Barber, L.K., Sweeney, R.W., Whitlock, R.H., Shulaw, W.P., Jacobson, R.H., Stabel, J.R., 2001. Evaluation of a commercial ELISA for diagnosis of paratuberculosis in cattle. J. Am. Vet. Med. Assoc. 218, 1163-1166.

Franken, P., 2005. Paratuberculosis in the Netherlands: the target and overview of activities. In: Manning EJB, Nielsen SS (Eds.), Proceedings of the 8th International Colloquium on Paratuberculosis. Madison, Wisconsin. pp 15-19.

Jubb, T.F., Sergeant, E.S.G., Callinan, A.P.L., Galvin, J.W., 2004. Estimate of sensitivity of an ELISA used to detect Johne's disease in Victorian dairy herds. Aust. Vet. J. 82, 569-573.

Kennedy, D.J., Allworth, M.B., Mitchell, R., 2002. Improving Australian assurance programs and risk management for Johne's disease. In: Juste RA (Eds.), Proceedings of the 7th International Colloquium on Paratuberculosis. Madison, Wisconsin. pp 424-428.

Kennedy, D.J., Benedictus, G., 2001. Control of *Mycobacterium avium* subsp *paratuberculosis* infection in agricultural species. Rev. Sci. Tech. 20, 151-179.

Kudahl, A.B., 2004. Economic consequences of paratuberculosis control in dairy herds. PhD Thesis. The Royal Veterinary and Agricultural University, Frederiksberg, Denmark.

Martin, P.A.J., Cameron, A.R., Greiner, M., 2007. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev. Vet. Med. 79, 71-97.

Martin, S.W., Shoukri, M., Thorburn, M.A., 1992. Evaluating the health status of herds based on tests applied to individuals. Prev. Vet. Med. 14, 33-43.

McKenna, S.L.B., O'Keefe, G.P., Barkema, H.W., Sockett, D.C., 2005. Evaluation of three ELISAs for *Mycobacterium avium* subsp. *paratuberculosis* using tissue and fecal culture as comparison standards. Vet. Microbiol. 110, 105-111.

Nielsen, S.S., 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subspecies *paratuberculosis in* dairy cattle. J. Vet. Med. B 49, 384-387.

Nielsen, S.S., Kolmos, B., Christoffersen, A.B., 2004. Comparison of contamination and growth of *Mycobacterium avium* subsp. *paratuberculosis* on two different media. J Applied Microbiol, 96, 149-153.

Nielsen, S.S., Grønbæk, C., Agger, J.F., Houe, H., 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53, 191-204.

Nielsen, S.S., Toft, N., 2006. Age-specific characteristics of ELISA and faecal culture for purpose specific testing for paratuberculosis. J Dairy Sci 89, 569-579.

Nielsen, S.S., Toft, N., Jorgensen, E., Bibby, B.M., 2007. Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response. Prev. Vet. Med. 81, 290-305.

Ridge, S.E., Morgan, I.R., Sockett, D.C., Collins, M.T., Condron, R.J., Skilbeck, N.W., Webber, J.J., 1991. Comparison of the Johne's absorbed EIA and the complement-fixation test for the diagnosis of Johne's disease in cattle. Aust. Vet. J. 68, 253-257.

Rogan, W.J., Gladen, B., 1978. Estimating prevalence from the results of a screening test. Am. J. Epidemiol. 107, 71-76.

Suess, E.A., Gardner, I.A., Johnson, W.O., 2002. Hierarchical Bayesian model for prevalence inferences and determination of a country's status for an animal pathogen. Prev. Vet. Med. 55, 155-171.

Sweeney, R.W., 1996. Transmission of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 12, 305-312.

Tavornpanich, S., Gardner, I.A., Carpenter, T.E., Johnson, W.O., Anderson, R.J., 2006. Evaluation of cost-effectiveness of targeted sampling methods for detection of *Mycobacterium avium* subsp. *paratuberculosis* infection in dairy herds. Am. J. Vet. Res. 67, 821-828.

Tavornpanich, S., Muñoz-Zanzi, C.A., Wells, S.J., Raizman, E.A., Carpenter, T.E., Johnson, W.O., Gardner, I.A., 2008. Simulation model for evaluation of testing strategies for detection of paratuberculosis in Midwestern US dairy herds, Prev. Vet. Med. 83, 65-82.

van Weering, H.J., Mars, M.H., Muskens, J., Middelesch, H., van Schaik, G., 2005. The effect of biosecurity measures for paratuberculosis on the seroprevalence in Dutch dairy herds. In: Manning EJB, Nielsen SS (Eds.), Proceedings of the 8th International Colloquium on Paratuberculosis. Madison, Wisconsin. pp 246-253.

Weber, M.F., Groenendaal, H., van Roermund, H.J., Nielen, M., 2004. Simulation of alternatives for the Dutch Johne's disease certification-and-monitoring program. Prev. Vet. Med. 62, 1-17.

Wells, S.J., Whitlock, R.H., Wagner, B.A., Collins, J., Garry, F., Hirst, H., Lawrence, J., Saville, W.J., Naugle, A.L., 2002. Sensitivity of test strategies used in the Voluntary Johne's Disease Herd Status Program for detection of *Mycobacterium paratuberculosis* infection in dairy cattle herds. J. Am. Vet. Med. Assoc. 220, 1053-7.

Whittington, R.J., Sergeant, E.S.G., 2001. Progress towards understanding the spread, detection and control of *Mycobacterium avium* subsp *paratuberculosis* in animal populations. Aust. Vet. J. 79, 267-278.

Paper IX

Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response

Nielsen SS, Toft N, Jørgensen E, Bibby BM, 2007

Published in
Preventive Veterinary Medicine 2007, 81, 290–305

Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response

S.S. Nielsen[1,*], N. Toft[1], E. Jørgensen[2], B.M. Bibby[3]

[1]Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark; [2] Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, University of Aarhus, P.O. Box 50, DK-8830, Tjele, Denmark; [3]Department of Biostatistics, University of Aarhus, Vennelyst Boulevard 6, DK-8000 Aarhus C, Denmark.

ABSTRACT

Diagnostic inference by use of assays such as ELISA is usually done by dichotomizing the optical density (OD)-values based on a predetermined cut-off. For paratuberculosis, a slowly developing infection in cattle and other ruminants, it is known that laboratory factors as well as animal specific covariates influence the OD-value, but while laboratory factors are adjusted for, the animal specific covariates are seldom utilized when establishing cut-offs. Furthermore, when dichotomizing an OD-value, information is lost. Considering the poor diagnostic performance of ELISAs for diagnosis of paratuberculosis, a framework for utilizing the continuous OD-values as well as known coavariates could be useful in addition to the traditional approaches, e.g. for estimating within-herd prevalences.

The objective of this study was to develop a Bayesian mixture model with two components describing the continuous OD response of infected and non-infected cows, while adjusting for known covariates. Based on this model, four different within-herd prevalence indicators were considered: the mean prevalence in the herd; the age adjusted prevalence of the herd for better between-herd comparisons; the rank of the age adjusted prevalence to better compare across time; and a threshold-based prevalence to describe differences between herds. For comparison, the within-herd prevalence and associated rank using a traditional dichotomization approach based on a single cut-off for an OD corrected for laboratory variation was estimated in a Bayesian model with priors for sensitivity and specificity.

The models were applied to the OD-values of a milk ELISA using samples from all lactating cows in 100 Danish dairy herds in three sampling rounds 13 months apart. The results of the comparison showed that including covariates in the mixture model reduced the uncertainty of the prevalence estimates compared to the cut-off based estimates. This allowed a more informative ranking of the herds where low ranking and high ranking herds were easier to identify.

1. INTRODUCTION

Paratuberculosis, also denoted Johne's disease, is an insidious, chronic infection of ruminants and other animals, caused by *Mycobacterium avium* subsp. *paratuberculosis* (Map). Due to the chronicity and slow development of infection, there is a profound lack of accurate diagnostic methods to correctly identify infected animals and thereby establish unbiased and precise estimates of the prevalence of infection at animal level.

Between the 1950s and 1980s, prevalence studies on cow-level were predominantly carried out using microbiological cultivation of tissues sampled at slaughter-plants.

Prevalence estimates from various countries and studies were in the range of 0–15% (summarised in Kennedy and Benedictus, 2001). Most studies used cultivation of samples from the ileo-cecal valve and/or lymph nodes to detect Map infection. While this method may be the most sensitive among available tests for detection of Map-infection, the sensitivity is still less than 70% (Whitlock et al., 1996). Furthermore, post-mortem examinations are expensive, time-consuming and are of limited value to farmers interested in the current prevalence of Map-infection in their herd. Selection bias may also be a prominent feature in the use of post-mortem samples.

In the 1990s, several prevalence studies were based on serological methods (summarised in Kennedy and Benedictus, 2001). Such studies, based on serology, e.g. ELISA, became popular, primarily because samples are easier to obtain, the test is cheaper and analyses can be automated. The outcome of an ELISA test is an optical density (OD), recorded on an ordinal scale. The OD-value is often dichotomised via a chosen threshold or cut-off. Apart from the infection status, covariates regarding laboratory factors, such as the effect of ELISA-plate, as well as animal traits, such as parity and stage of lactation, all influence the OD response (Nielsen et al., 2002a). To adjust for the laboratory factors, the OD is often corrected to account for the variation between ELISA-plates, while the latter is traditionally ignored to facilitate the use of only one cut-off of the (corrected) OD. However, in previous studies, these animal specific covariates have been demonstrated to have a statistically significant influence on the OD-value (Nielsen et al., 2002a; Toft et al., 2005a). In particular age can be expected to influence the OD-value, due to the chronic nature of the infection. Therefore, age related characteristics and other covariates should be considered when interpreting dichotomised ELISA-results, i.e. the cut-off value should depend on such covariates.

Generally, the underlying assumption of any test scenario is that there are two groups of animals (infected and non-infected) which differ in test response. For culture based methods, this difference is characterised by growth versus no growth, with some degree of misclassification. For continuous tests, such as ELISAs, the difference is in the average response of infected versus non-infected, potentially with a difference in variation around these group means. Hence, when applied to a population of animals with unknown infection status, the observed OD-values will be a mixture of the distributions of ODs from infected and non-infected animals in the proportion determined by the prevalence in the sampled population. The parameters of the underlying mixture model, i.e. the proportion of the mixtures and the mean and variance of the OD distributions including the appropriate covariates, can be estimated from the data. However, depending on the nature of the problem, additional assumptions might be required. Therefore, a flexible approach to mixture modelling is to use Bayesian analysis, where assumptions are added as prior distributions reflecting the uncertainty associated with the prior knowledge.

The objective of our study was to develop an alternative method to estimate the within-herd prevalence of paratuberculosis in 100 Danish dairy herds using the continuous OD response of a milk ELISA adjusted for age covariates and laboratory effects through the use of a Bayesian mixture model using Markov Chain Monte Carlo (MCMC) based methods, i.e. WinBUGS (Spiegelhalter et al., 2004). Based on the formulated model, several measures of the occurrence of infection at herd level can be calculated. We present and discuss these with respect to their interpretation and potential application in the control of paratuberculosis at the animal, herd and regional level. For comparison, a traditional cut-off based approach

to the interpretation of the OD response of an ELISA was formulated in a Bayesian framework to estimate the prevalence of paratuberculosis using prior obtained information about the sensititivy and specificity of the ELISA at the selected cut-off in a comparable population.

2. MATERIALS AND METHODS

2.1. Area, herds and samplings

The sample region was defined by four postal zip codes in Southern Jutland, Denmark, where approximately 280 dairy herds were located in 1999 (Andersen et al., 2000), and of which 110 farmers agreed to participate in projects relating to paratuberculosis. Three sample rounds were performed, with approximately 13 months between each sample in the individual herd. Of the original 110 herds, two withdrew from the study after the first sampling period, four herds stopped production between the first and second sampling period and four more herds stopped production between the second and third sampling period. Thus, the first sample was from 110 herds; the second was from 104 herds and the third was from 100 herds. Only the 100 herds which contributed to all three samples were included in the present study. The first sampling period was between August 1999 and February 2000; the second sampling period was from September 2000 to March 2001; and the last sampling period was from October 2001 to April 2002. A brief summary of herd demographics is given in Table 1. The number of cows represents the number of lactating cows in the herd on the test-date, though the herd size also included dry cows (average proportion of dry cows in the herds was 15%). There were no formal control programs initiated during the study period, however, herd owners might have implemented individual control measures.

2.2. Cows and samples

Milk samples were obtained from all lactating cows through the milk yield scheme at each sampling. Subsequently, information about breed, parity, age at 1st calving (AC1) and days in milk (DIM) for each cow in the herd at the time of sampling, was retrieved from the Danish Cattle Database. In Table 1, the distribution of breeds, parity (divided in 1^{st}, 2^{nd} and $>2^{nd}$), AC1 and DIM is given for each of the three sample rounds. A total of 21,912 samples were collected from 14,343 cows, with 8506 cows being sampled once, 4107 cows twice, 1728 cows three times and 2 cows four times. The latter two had been traded between the herds between sampling of the individual herds.

2.3. Detection of antibodies

Antibodies present in the milk samples were detected using an ELISA described in detail in Nielsen (2002). The level of antibodies was estimated by the recording the OD-values; it should be noted that the recordings were kept on a continuous scale. The ELISA-plate identification number was recorded to enable correction for laboratory variation. The number of ELISA-plates used in each sampling period and a description of the raw OD-values recorded is given in Table 1.

Table 1. Descriptive statistics of herds, cows (breed and age-information) and ELISA-parameters divided by sampling period in 100 sampled Danish dairy herds from 1999-2003

	Sampling period		
	1	2	3
Parity			
1	2700	2866	3067
2	1908	1848	2048
>2	2504	2477	2494
Herd size (no. of cows on test date)			
Min	19	16	21
Q1	53	50	57
Median	62	67	66
Q3	85	88	91
Max	210	205	236
Breeds (no. of cows on test date)			
Red Danish	164	145	150
Danish Holstein (Black)	5796	5846	6239
Danish Jersey	746	791	818
Danish Holstein (Red)	16	16	20
Crossbreed	384	382	372
Other	6	11	10
Age 1st calving (days)			
Min	490	460	495
Q1	780	782	782
Median	835	836	835
Q3	907	909	907
Max	2000	2310	2000
Days in milk (DIM) (days)			
Min	7	7	7
Q1	77	79	82
Median	160	160	165
Q3	251	251	254
Max	400	400	400
No. ELISA plates			
	193	209	215
Raw OD-values			
Min	0.003	0.005	0.008
Q1	0.048	0.051	0.055
Median	0.075	0.089	0.093
Q3	0.127	0.165	0.169
Max.	2.252	2.460	2.498

2.4. Statistical model

A finite mixture model (Gelman et al., 1995) with two mixture components representing infected and non-infected dairy cows, respectively, was used in this study. Previous studies of the milk ELISA test (Toft et al., 2005a) showed that the log-transformed ODs were reasonably approximated by a normal distribution within each of the infection groups, when accounting for additional covariates. Furthermore, the above study demonstrated an effect of age by means of parity and DIM on the mean log(OD) response. The latter association was

not linear (on the log scale), and as a result cows were grouped into three groups: 0–1 weeks after calving; 2–28 weeks after calving; and >28 weeks after calving. Additionally, cows were grouped into first, second and above second parity to avoid too few observations for older cows. To account for the variation between plates, a random effect of the ELISA-plate was assumed, with separate mean and variance for each sampling period. Finally, the variance was allowed to vary between the infected and non-infected group to reflect that more variation in ELISA response was expected in the infected group. Conditioned on the (latent) infection status (Z) of the sampled animal, the model assumed that the log-transformed OD-values followed a Normal distribution, where the mean and variance were controlled by the infection status and the covariates discussed above. The (latent) infection status (Z) was either infected or not infected, i.e. Z was Bernoulli distributed, with parameter p defining the probability that the log(OD) was from a truly infected animal. This probability was expected to be influenced by parity and stage of lactation due to the nature of paratuberculosis. Furthermore, it was assumed that there is an effect of herd and sampling period on the probability of infection to allow herd mates to be more similar than non-herd mates. Therefore, the following finite mixture model with two mixture components for the log-transformed OD was specified:

$$\log(OD)_i \mid Z_i \sim N(\mu_i; \sigma^2_{Z_i})$$
$$Z_i \sim Bern(p_i)$$
$$\mu_i = \alpha_{Z_i Parity_i} + \beta_{Z_i DIM_i} + C_{Plate_i}$$
$$C_{Plate} \sim N(\eta_{Period_i}, \sigma^2_{Period_i})$$

$$\text{logit}(p_i) = \delta_{Parity_i} + \kappa_{Period_i Herd_i}$$
$$\kappa_{Period\,Herd} = \begin{cases} \gamma_{Herd_i} & Period = 1 \\ \gamma_{Herd_i} + \varepsilon_{2,Herd_i} & Period = 2 \\ \gamma_{Herd_i} + \varepsilon_{3,Herd_i} & Period = 3 \end{cases}$$
$$\gamma_{Herd} \sim N(0, \sigma^2_\gamma)$$
$$\varepsilon_{2,Herd}, \varepsilon_{3,Herd} \sim N(0, \sigma^2_\varepsilon)$$

In this model,

$\log(OD)_i$ is the observed log-transformed OD value for the *i*th sample which conditional on the Bernoulli distributed infection status (Z_i) follows a normal distribution with mean μ_i and variance $\sigma^2_{Z_i}$;

$\alpha_{Z_i Parity_i}$ is the combined effect of the (latent) infection status and parity on the mean log(OD).

$\beta_{Z_i DIM_i}$ is the combined effect of the (latent) infection status and stage in lactation on the mean log(OD);

C_{Plate_i} is the random effect of plate within each round, with varying mean to model potential changes in the average log(OD) response between sampling periods;

p_i is the probability that the *i*th sample is from an infected cow;

δ_{Parity_i} is the effect of parity on the probability of infection for the *i*th sample;

and $\kappa_{Period_i Herd_i}$ is the combined effect of the round and herd on the probability of infection modeled as a random effect of herd (γ_{Herd}) within the first period and an additive random effect of herd within each of the following two periods ($\varepsilon_{2,Herd}, \varepsilon_{3,Herd}$).

The α's were restricted so that for each parity, the mean log(OD) of the non-infected cows was strictly smaller than the mean log(OD) of the infected cows. This was done in order to address the implicit unidentifiability of any mixture model (the problem of label switching, i.e. is group one the infected or the non-infected population?); and to avoid problems with convergence due to the MCMC sampler getting stuck in the trivial solution of all p_i = 1 (or 0).

Parameters were generally given only vaguely informative priors allowing variation within the relevant range, but disallowing extreme values. For example, the priors on α′s representing the effect of parity on the non-infected group were given uniform priors (Uniform (−4,−1)), which allow variation within the relevant range, but prevents extreme values. The α′s representing the effect of the infected group, were then specified as the corresponding α for the non-infected, but with a strictly positive contribution added as a Gamma distribution (Gamma (5,3) with mean (5/3 = 1.67) and standard deviation ($\sqrt{5/9}$ = 0.75)), which proved to be neither too restrictive nor too relaxing to ensure convergence and not influencing the posterior distributions in any particular way. The prior distribution for all variances were the same, Uniform (0.001,5), the prior distribution for the β′s was Uniform (−1,0.5), however, the β′s were modelled as an additive effect on α, therefore most were restricted to 0. Finally, the prior for the mean effect of plate (η_{Period}) was also additive, so that the mean for sampling period 1 was fixed at 0, while the mean for periods 2 and 3 were given prior distributions $N(0,1)$.

Model selection was done by calculating the mean deviance for the candidate models. For finite mixture models the number of actual parameters in a model is not well-defined. Therefore, the magnitude of changes in deviance was observed, but no formal testing was done. The Deviance Information Criteria (DIC) (Spiegelhalter et al., 2004 D.J. Spiegelhalter, A. Thomas and N.G. Best, WinBUGS Version 1.4 User Manual, MRC Biostatistics Unit, Cambridge, United Kingdom (2004).Spiegelhalter et al., 2004), which is an alternative for model selection in Bayesian models, was not possible to calculate due to the problem with the number of parameters being undefined. Model checking was carried out by inspecting the Bayesian residuals against the expected values and by calculating the probability of getting more extreme observations (Spiegelhalter et al., 1995). As a goodness-of-fit measure, Bayesian p-values (Spiegelhalter et al., 1995) were calculated to test whether the skewness and the kurtosis of the observed data were consistent with the assumed normal distribution when compared to simulated data from a normal distribution.

2.5. Herd prevalences

Using the above described model, it was possible to derive several different measures reflecting the prevalence of Map-infected cows. We present four of them.

2.5.1. Mean prevalence: As p_i gives the probability of infection for the *i*th cow, the herd specific prevalence for each period could be calculated directly as the mean of the relevant individual p_i's,; this measure will be referred to as the mean prevalence (*mp*).

2.5.2. Age adjusted prevalence: While *mp* gives the direct prevalence estimated using the model, it might be of more interest to obtain an estimate where the effect of parity has been accounted for, therefore allowing comparison of herd specific prevalences at a given common age structure. The prevalence estimates were calculated for an age distribution of 40, 25 and 35% for first parity, second parity and older cows respectively in order to reflect the typical age structure in a Danish dairy herd. Hence the age adjusted prevalence (*ap*) for the *h*th herd in the *r*th period is given by

$$ap_{rh} = 0.40\,p_{Parity=1;rh} + 0.25\,p_{Parity=2;rh} + 0.35\,p_{Parity>2;rh}$$

where e.g. $p_{Parity=1;rh}$ is the prevalence in the *h*th herd in the *r*th period, i.e.

$$\text{logit}(p_{Parity=1;rh}) = \delta_1 + \kappa_{rh}$$

This value as well as the other two was calculated in WinBUGS using the individual samples, rather than mean values of the parameters.

2.5.3. Rank of age adjusted prevalence: Rather than using the age adjusted prevalence estimates directly, it might be preferred to use the rank (*rp*) of the age adjusted prevalence for a herd among the 100 herds within each round, thus eliminating potential drifts in the OD across time when comparing sampling periods. The higher the rank, the higher the prevalence, i.e. rank 1 reflects the lowest age adjusted prevalence.

2.5.4. Individual threshold prevalence: The mean prevalence and the age adjusted prevalence might cover large differences between individual animals, i.e. a herd prevalence of 50% (in a small herd with two cows) might be from two animals each with 50% probability of infection or 10% and 90% probability, respectively. Thus, we define the individual threshold prevalence (*ip*) as the average proportion of individuals in the herd whose probability of infection lies above a certain threshold, e.g. 50%.

2.6. Cut-off based prevalence

To compare the prevalence estimates from the above model to a more traditional application of an ELISA for prevalence estimation, we used data from a previous study (Nielsen et al., 2002b) to estimate the sensitivity (Se) and specificity (Sp) for a dichotomized ELISA-result using a fixed cut-off for the corrected OD (an OD-value corrected for variation between ELISA-plates, using the method described in Nielsen et al. (2002a)). Using a Bayesian version of the latent class model by Hui and Walter (1980) for estimating properties of diagnostic tests in the absence of a gold standard (Toft et al., 2005b), the posterior mean (m) and standard deviation (S.D.) for Se and Sp of the ELISA at the cut-off mostly used (ELISA-positive >0.3 for the corrected OD-value) were estimated as mSe = 0.5101, S.D.Se = 0.0803 and mSp = 0.9289, S.D.Sp = 0.009376, respectively. These values were then used to form prior distributions for Se and Sp as Beta distributions: Se Beta(19.26,18.50) and Sp Beta(696.94,53.35). Using these priors and a non-informative Beta(1,1) for the 'true' prevalence (cp) the usual association between the count of test positives (r) and the number of animals tested (n) were assumed, i.e. r Binomial(sp,n), where sp is the sero-prevalence or apparent prevalence, defined as sp = Se × cp + (1 − Sp) × (1 − cp). Although cp traditionally is referred to as the 'true' prevalence, we will refer to it as the classic prevalence to reduce potential reader bias inferred by the term 'true'. Furthermore, analogous to (rp), we defined (rc) as the rank of cp among the 100 herds within a sampling period.

2.7. Model implementation

The mixture model and the classic prevalence model were both implemented in WinBUGS (Spiegelhalter et al., 2004). For each model, the first 5000 iterations were discarded as a burn-in and the following 10,000 iterations were kept for posterior inference. Convergence of the chain after the initial burn-in was assessed by visual inspection of the time-series plots for selected variables, i.e. all hyperparameters and a subset of parameters for the individual animals and farms in the mixture model, as well as Gelman–Rubin diagnostic plots using three sample chains with different starting values (Brooks and Gelman, 1998). For a discussion of the convergence of MCMC chain and tools for assessing convergence, see (Toft et al., 2007).

3. RESULTS

Histograms of the log-transformed OD-values are given in Fig. 1, stratified by parity. The estimated posterior mean log(OD) for the infected and non-infected groups (for round 1) for each parity are superimposed (see Table 2), with the infected group being adjusted for the effect of DIM corresponding to the 2–28 weeks group, which represents the largest of the three groups. The horizontal lines represent the range covering 95% of the probability mass for each of the two sub-populations.

In Table 2, the posterior distributions of the parameters for the mixture model are summarized using the estimated posterior means and 2.5%- and 97.5%-percentiles. There appears to be an almost constant difference of approximately one log(OD) between the infected and non-infected group within each parity (α). The effect of stage of lactation (β) is primarily seen as a reduction in log(OD) response for the mid-lactation cows compared to early or late stage lactating cows. The effect of ELISA-plate (σ^2_{Period}) is also substantial and there is more variation between ELISA-plates for sampling period two compared to sampling period one and three. The 95% credibility posterior intervals (CPI) for the effect of parity are markedly narrower for the non-infected group than the infected group.

The estimates for the effect of parity on the probability of infection (δ) showed the expected lower probability of infection among first parity cows compared to second parity and older cows, but there also appeared to be a higher probability of infection for second parity cows compared to older cows, suggesting that there is reduced antibody response when time from infection increases. The variance component of the effect of herd and the additive effect of herd within sampling period two and three appear to be quite large, suggesting that there is substantial variation between herds as well as within herds between periods.

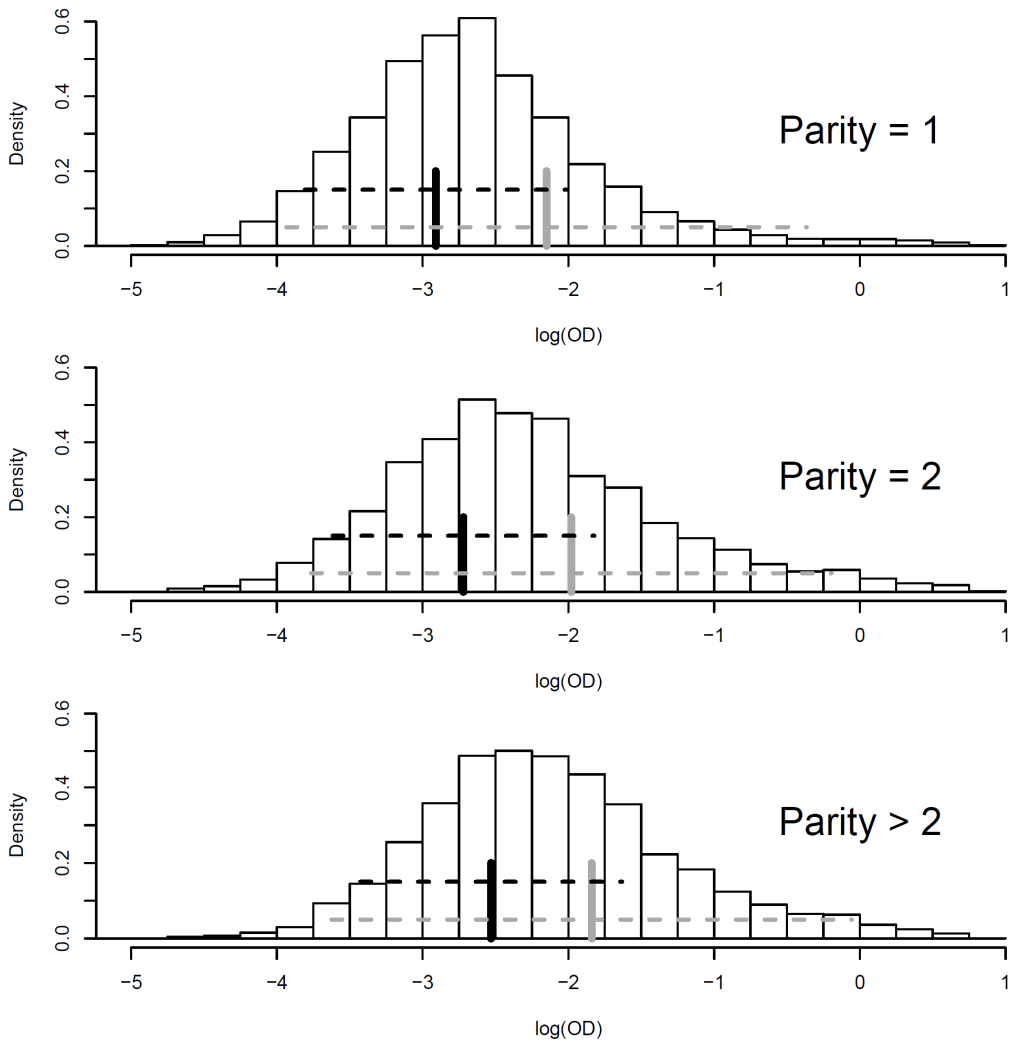


Figure 1. The distribution of log(OD) stratified by parity, the estimated mean posterior log(OD) (see Table 2) for the infected (in mid-lactation) (grey vertical bar) and non-infected sub-population (black vertical bar) have been imposed on the histograms for each parity. The horizontal dashed lines represent the range in which 95% of the probability mass is located for each of the two sub-populations

Table 2. The summary estimates of the posterior distributions of the parameters for the finite Bayesian mixture model on within-herd prevalence estimates for bovine paratuberculosis

				95% CPI[§]	
	Model component	Parameter	Mean	$Q_{0.025}$	$Q_{0.975}$
log(OD)					
Fixed effects					
	Infection x Parity	$\alpha_{Z=Yes,\ Parity=1}$	-1.81	-1.98	-1.64
		$\alpha_{Z=No,\ Parity=1}$	-2.91	-2.97	-2.85
		$\alpha_{Z=Yes,\ Parity=2}$	-1.64	-1.79	-1.48
		$\alpha_{Z=No,\ Parity=2}$	-2.72	-2.78	-2.66
		$\alpha_{Z=Yes,\ Parity>2}$	-1.50	-1.65	-1.34
		$\alpha_{Z=No,\ Parity>2}$	-2.53	-2.60	-2.47
	Infection x DIM	$\beta_{Z=Yes,\ DIM=0\text{-}1\ weeks}$	0	.	.
		$\beta_{Z=Yes,\ DIM=2\text{-}28\ weeks}$	-0.34	-0.48	-0.19
		$\beta_{Z=Yes,\ DIM>28\ weeks}$	-0.02	-0.17	0.12
	Plate x Period	$\eta_{Period=1}$	0	.	.
		$\eta_{Period=2}$	0.11	0.02	0.20
		$\eta_{Period=3}$	0.16	0.08	0.24
Random effects					
	Plate x Period	$\sigma^2_{Period=1}$	0.15	0.12	0.19
		$\sigma^2_{Period=2}$	0.25	0.20	0.30
		$\sigma^2_{Period=3}$	0.13	0.11	0.16
Variance					
	Infected	$\sigma^2_{Z=Yes}$	0.84	0.80	0.88
	Non-infected	$\sigma^2_{Z=No}$	0.21	0.20	0.22
Probability					
Fixed effects					
	Parity	$\delta_{Parity=1}$	-1.87	-2.16	-1.57
		$\delta_{Parity=2}$	-0.66	-0.93	-0.41
		$\delta_{Parity>2}$	-0.77	-1.02	-0.51
Random effects					
	Herd effect	σ^2_{γ}	0.89	0.58	1.29
	Period effect	σ^2_{ε}	0.93	0.61	1.32

[§]95% CPI: 95% credibility posterior interval

To assess the fit of the above model, three measures were used. First, different models, i.e. variations in covariates and formulation, were compared using the mean deviance; the current model was generally favoured (results not shown). Specifically, the model with two mixtures (i.e. allowing infected as well as non-infected animals in the population) was preferred over a model with only one component (i.e. a model where all animals had the same infection status), and the models allowing variation between rounds (in both prevalence and mean OD response) were preferred to models without this variation.

Second, the probability of observing a more extreme observation was calculated for each individual log(OD) using the current model. These probabilities were then plotted against the log(OD), producing a bell-shaped curve (results not shown), in which there appeared to be some distortion for the log(OD)s in the interval between the average log(OD) of the infected and non-infected groups, respectively. This again suggests a lack of diagnostic sensitivity of the milk ELISA, which gives rise to problems when trying to separate the infected and non-infected groups, as also illustrated by the overlap of the two 95% probability ranges for the infected and non-infected sub-groups imposed in Fig. 1. Finally, the model fit of the present model was assessed by comparing the skewness and kurtosis of the observed values (conditional on a simulated infection status) to those expected from a normal distribution (conditional on infection status), i.e. skewness = 0 and kurtosis = 3. The mean posterior skewness was 0.07 with a Bayesian p-value of 0.03 (calculated as the proportion of iterations where the observed skewness was larger than that of a simulated dataset). Therefore, it appears that the observed data is a little more asymmetric than those from a normal distribution. The observed mean posterior estimated kurtosis was 3.06, with a Bayesian p-value of 0.25, suggesting that the data are neither peaked nor flat compared to those from a simulated normal distribution.

To compare the model-based estimates to the classic approach, the posterior mean and associated 95% CPI for ap and cp are shown for sampling period one in Fig. 2 as well as the median ranks and 95% CPI for the 100 herds based on these estimated ap and cp, i.e. the median and 95% CPI of rp and rc. The results are similar for sampling round two and three (data not shown).

To summarize our findings for the entire population, the mp, ip, and apparent prevalence based on the cut-off used to calculate cp were calculated for the population stratified by round and parity (1, 2 and >2) and presented in Table 3. These estimates give the appearance of a general increase in the prevalence of paratuberculosis between the first and second sampling period. However, an increase in prevalence of this magnitude does not seem realistic and might indicate that there is a shift in the overall ELISA response between the first and subsequent sampling periods. It is not possible to confirm this within the current study, but does question whether the prevalence estimates can be compared across rounds, as a result favouring the rank estimate (rp) when discussing changes through time. This is ignored in Fig. 3, where the empirical distributions of the mean posterior individual mp and ip for the entire sample population aggregated across sample rounds are given.

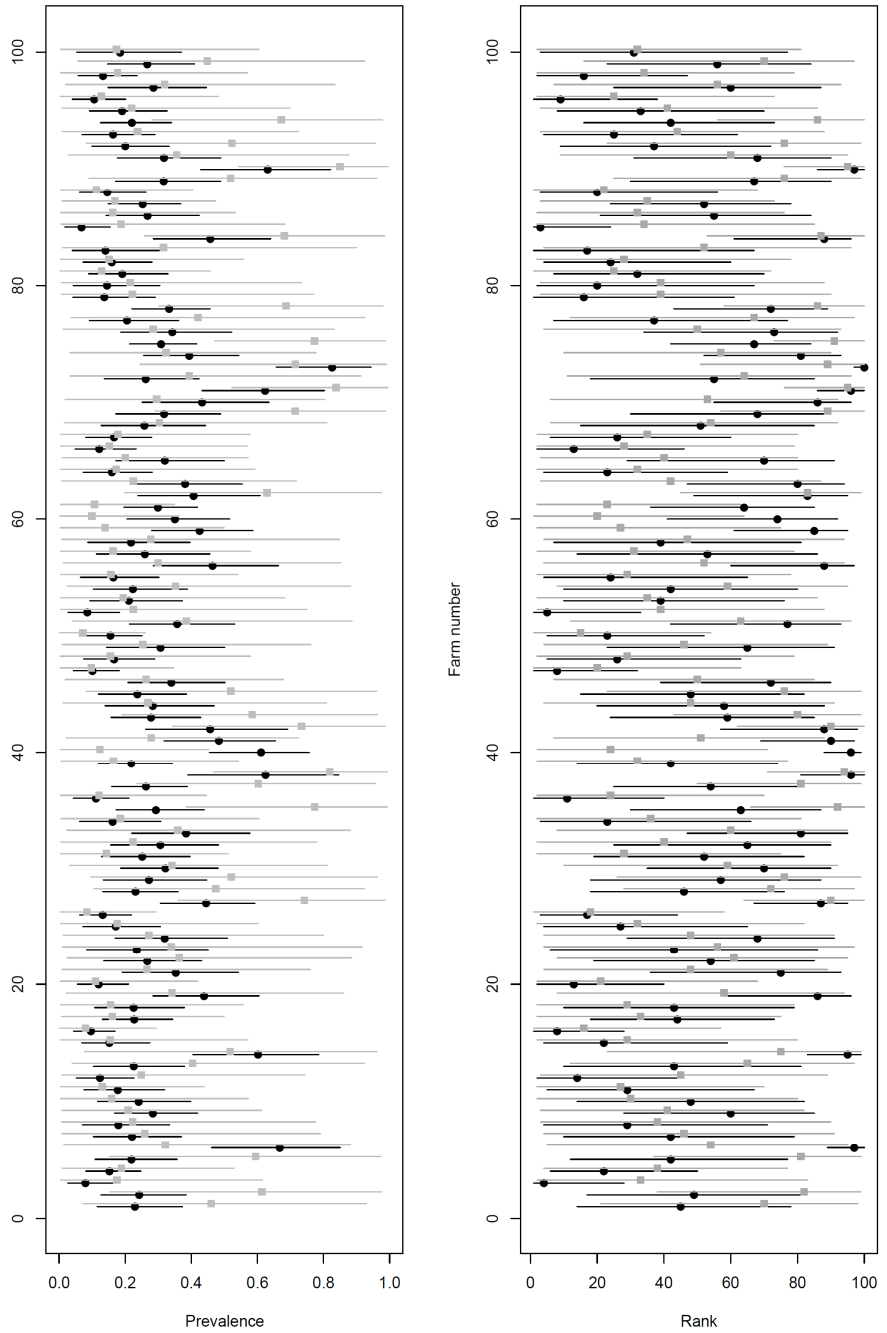


Figure 2. Left side: The mean age adjusted prevalence (*ap*) (black circle) and classic prevalence (*cp*) (grey square) with associated 95% CPI for 100 herds in sampling period one. Right side: The median rank of age adjusted prevalence (*rp*) (black circle) and rank of classic prevalence (*rc*) (grey square) and the associated 95% CPI for 100 herds in sampling period one. Estimates were obtained in a Bayesian model for within-herd prevalence estimation of bovine paratuberculosis.

Table 3. The estimated posterior mean of the mean (within-herd) prevalence (mp),, individual threshold prevalence (*ip*) and the empirical mean of the cut-off based apparent prevalence ($AP_{cut\text{-}off}$) for the individual cows, stratified by sampling period and parity, estimated in a Bayesian model on within-herd prevalence estimates for bovine paratuberculosis

	Period 1			Period 2			Period 3		
Parity	1	2	>2	1	2	>2	1	2	>2
mp	16	35	33	22	42	39	21	42	39
Ip	3	18	15	7	36	30	7	35	28
$AP_{cut\text{-}off}$	2	4	4	3	7	8	4	8	10

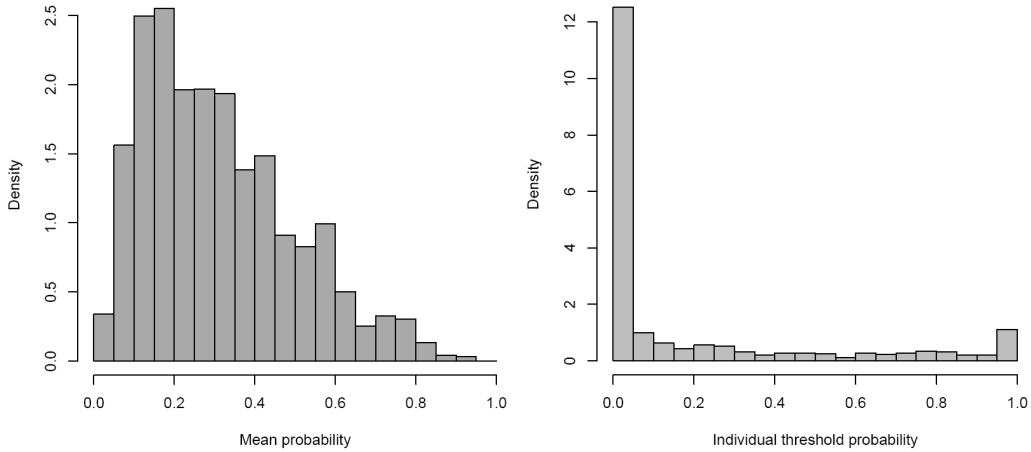


Figure 3. The empirical distribution of the mean posterior estimates for the individual cows of a) the mean probability (mp) of infection for the individual animals; b) the probability that mp exceeds 50% for the individual animals (ip) in a Bayesian model for within-herd prevalence estimates of bovine paratuberculosis.

Table 4. The Spearman correlation between the six different prevalence estimates, i.e. the four model-based estimates (m*p, ap, ip, rp*) and the two cut-off-based prevalence estimates (*cp, rc*) calculated using data from the first sample round only, estimated in a Bayesian model on within-herd prevalence estimates for bovine paratuberculosis

	mp	*ap*	*ip*	*rp*	*cp*	*rc*
mp	1	0.99	0.93	0.93	0.55	0.54
ap		1	0.93	0.93	0.55	0.64
ip			1	0.78	0.46	0.43
rp				1	0.55	0.58
cp					1	0.98
rc						1

Since all prevalence estimates are measuring the same thing, they would naturally be expected to be strongly correlated. In Table 4 the correlations between the six different prevalence estimates are shown, based on the samples from the first period only. The correlations based on data from the other periods are similar to these. As expected there is a high correlation among the four model-based estimates and the two cut-off based estimates, but less correlation between model-based and cut-off based estimates.

4. DISCUSSION

Using the continuous OD response of a milk ELISA, we estimated the within herd prevalence of Map infected cows in 100 Danish dairy herds using a Bayesian mixture model. The model includes known covariates: animal factors, such as parity and stage of lactation, and laboratory effects, i.e. variation between ELISA-plates. Using this model, four different measures for the occurrence of Map-infection in herds were estimated, i.e. *mp*, *ap*, *rp* and *ip*. For comparison, a traditional cut-off based estimate of Map prevalence (*cp*), and its rank (*rc*), were also calculated using ODs corrected for the effect of ELISA-plate and adjusted for misclassification using the Se and Sp of the ELISA at the given cut-off, but ignoring animal covariates, i.e. the same cut-off value were used for all cows.

The comparison of *ap* and *cp*, as well as the associated *rp* and *rc* in Fig. 2 clearly shows the impact of the covariates used for the model-based prevalence measures; the 95% CPI are considerably narrower for *ap* and *rp*, when compared to *cp* and *rc*. These covariates have a statistical significance, to the extent that the current model was favoured over the alternatives it was compared against. However, whether this increase in the precision of the estimates has any influence on the decisions is an open question that will be left for future studies. Our concern was to devise a method which could utilize covariates and information from other herds to improve the estimate of prevalence in individual herds. It is apparent that the narrower width of the 95% CPI is an improvement, the *rp* seems to be more useful in defining a group of low ranked herds, high ranked herds and a middle group, than the corresponding *rc*, which reflects the ranking when only the test results are used.

Our model is a Bayesian model and relies on information other than that obtained from data; specifically, we imposed restrictions on the OD response of infected versus uninfected cows, by enforcing that the expected log(OD) of an infected cow with a certain parity and stage of lactation was strictly greater than that of a similar uninfected cow. A claim that is justified by existing literature to a great extent, e.g. it forms the basis of most test evaluations of ELISAs. While similar associations could be imposed on the effect of parity and stage of laction, we chose not to do so, due to the large quantity of data available. The strength of the Bayesian framework is that the uncertainty about the prior knowledge can also be represented, since prior information is specified as distributions. This is why we chose to compare our model to a traditional application of an ELISA for prevalence estimation in a Bayesian formulation. To calculate the (true) prevalence based on the observed sero-prevalence, prior information about the sensitivity and specificity is also needed for the traditional approach. This information can also be obtained from the literature. Here we used data from a previous research project where the milk ELISA was evaluated in a latent class model against a faecal culture test (Nielsen et al., 2002b), however, the source of the information is not important. It is more important to acknowledge that Se and Sp of a test cannot be considered as point estimates, but have an associated uncertainty, which is easily addressed using a Bayesian model with the Se and Sp specified as priors. The priors we used should reflect the best available information about the milk ELISA when applied in this specific area, since the above study was conducted in the same area as the current study. Despite this, when applied to our data the posterior 95% CPI were Se:[0.10,0.12] and Sp:[0.98,0.99] which reflects the given priors rather poorly and suggests that the ELISA behaves differently when applied to our data. This should be taken into account when comparing the traditional approach to our model-based approach. There appeared to be

much less need for prior information and it could be given in a rather vague form, whereas the traditional model needs informative priors to give meaningful estimates of prevalence.

Regarding the different model-based prevalence estimates, there was little difference between mp and ap, especially when considering the uncertainty associated with the estimates. This indicates that either the (significant) effect of parity does not change the estimate of prevalence in a biologically important manner or there is a relatively stable age distribution across the 100 herds. The interpretation of the estimates is essentially the same: they are the mean of the estimated individual probabilities of Map-infection for the tested cows in a herd (or the population related to the estimate). As such it is probably the best unbiased estimate of the prevalence in a herd that we are able to get from a milk ELISA, when prevalence is interpreted as the proportion of Map-infected individuals. However, it appears that no herd is Map-free when using the mean or age adjusted prevalence. Whether this is correct is questionable as illustrated in Fig. 1, where it is evident that even for small values of log(OD) we cannot decisively say that it is the result of a truly non-infected cow. This suggests that there is a potential need for future research into a model where a mixture component at the herd level is added to divide herds into infected and non-infected.

A probability of infection of about 10% for an individual cow provides little evidence about the cow being infected, in fact it is evidence of the opposite, since we are 90% sure that it is not infected. But is there a difference between an estimated mean prevalence of 9% (herd 3, period 1) and 22% (herd 1, period 1)? A probability of infection at 22% for an individual cow is still more indicative of the cow being non-infected, than infected. Therefore, for some applications, there might be other summary measures for a herd which are more relevant. Comparing herd 19, period 1 (*mp* = 48%) and herd 21, round 1 (*mp* = 35%), there is again a difference of about 13%-point, but in herd 19 approximately 52% of the individual cows have an estimated probability of infection above 50% (*ip*), compared to only *ip* = 18% in herd 21. This difference may give rise to different approaches towards reducing the prevalence in the herds.

To some extent, it might be more relevant to compare herds to other herds without too much emphasis on the actual level of infection, e.g. use the rank of the prevalence rather than the prevalence itself. This is partly reinforced by the unexpected, and to some extent, unrealistic increase in prevalence between sampling period one and two, where a possible explanation could be a change in the performance of the ELISA. The ELISA-protocol has not changed, but it is not uncommon that ELISA tests are sensitive to changes in laboratory factors, e.g. room temperature, laboratory technician or other unknown factors. This should have been accounted for by including the effect of ELISA-plate as a combination of fixed and random effects, but apparently not enough. Comparing estimates in Table 3, we note that the change has also not been reflected by the correction of the OD, since the same unexplainable increase in prevalence is seen in the cut-off based apparent prevalence. Due to these unforeseen problems, comparisons across the sample rounds using the rank prevalence (*rp*) are probably better justified. However, to some extent this is more informative than the prevalence itself. When a herd participates in a control program, it is expected that the prevalence of Map-infected cows declines over time, and although this decline itself is relevant, it is just as relevant to benchmark the prevalence against other herds in a control program, to determine how successful the herd manager is in implementing the control mechanisms. For this purpose, it is more informative to know if ones rank among the participating farms has changed, rather than the exact prevalence. The

100 herds used in this study have not been subject to any official control program; hence changes in rank between the sample rounds must be attributed to the natural variation and perhaps the individual control strategies which some herd-managers might have adopted during the study period. From Fig. 2, it is clear, that some measure of uncertainty must be used when communicating the rank of a herd to the owner, but it is also apparent that the use of covariates and information from other herds reduces the uncertainty about the rank or prevalence of a herd.

Control of paratuberculosis is hampered by the lack of accurate diagnostic tests. The ELISA used in this study has been evaluated as a dichotomous test with fairly discouraging results (Nielsen et al., 2002b), but there is no evidence that it is better or worse than most available tests (e.g. Collins et al., 2005; McKenna et al., 2005). Table 2 clearly illustrates the problems associated with the milk ELISA; a relatively constant difference of approximately 1–1.1 in the average log(OD) response between the infected and non-infected groups stratified by parity. However, there is a large variation within the infected group, whose variance is about four times larger than the variance of the non-infected group. This is expected due to the slow developing nature of the Map-infection, but it also adds support to the belief that paratuberculosis is better represented as a multi-stage disease (Toft et al., 2005). Comparing the difference between infected and non-infected, to the contributions of the other covariates affecting the mean log(OD) response, it is apparent that these covariates cannot be ignored without a serious loss of discriminatory power of the ELISA, e.g. the contribution of including DIM as a covariate represents a change in mean log(OD) corresponding to about 30% of the difference between infected and non-infected.

Our model fit might need improvement, since the assumption of the log(OD) following a normal distribution conditional on infection status was not supported by the test of skewness of the observed data compared to data from a simulated distribution. However, it is unlikely that another distribution would give a better fit unless the flexibility of such a distribution would make it even harder to discriminate between the infected and non-infected cows. It is indisputable that there is a difference in the log(OD) response of infected and non-infected cows, since this has been established elsewhere by comparing the ELISA response of culture positive cows and cows from negative herds (Nielsen et al., 2002c). Our definition of infected does not necessarily include only those cows shedding Map, but in a latent class model, such as the one used in the current study there is no way of formally defining what constitutes an infected cow, since this to some extent depends on the chosen test(s). It is worth to note, that latent class analysis in recent years has been established in the literature as an accepted alternative to traditional test evaluations, and is recognized by the OIE.

The framework we have presented here can be extended to accommodate more tests, if e.g. results from faecal culturing of some or all cows were available, these results could be incorporated in the model to strengthen the estimation of the probability of infection for the individual cows. Similarly, repeated testing of individual cows could be utilized. We have ignored any individual cow-effect, i.e. potential correlation between samples from the same cow between rounds. Conditional on infection status the cow-effect did not appear to be very large, and excluding this effect reduced the complexity of the model, which helped to reduce the computation time.

Even in its present form, the model requires a computation time of about 10 h for 15,000 iterations of one chain, i.e. about 30 h when running the model with three starting points to assess convergence.

5. CONCLUSION

The herd prevalence of Map-infected cows can be estimated using the continuous OD response of an ELISA in a model, allowing covariates of laboratory factors as well as cow-specific information, thus improving the diagnostic power of the ELISA. We have demonstrated that dichotomizing the test result of an ELISA is unnecessary and to some extent counter-productive. Furthermore, we have demonstrated that information from herds may be combined when estimating the prevalence of the individual herds. Finally, we showed that from the model several measures of disease occurrence could be obtained, each giving the decision maker potentially useful information.

ACKNOWLEDGEMENTS

The Danish Cattle Federation financed this study, partly via Kongeåprojektet, partly via CEPROS-II/7. Nils Toft was financed by a grant from the Danish Research Council.

REFERENCES

Brooks, S.P., Gelman, A., 1998. Alternative methods for monitoring convergence of iterative simulations. J. Comp. Graph. Stat., 7, 434-455.

Collins, M.T., Wells, S.J., Petrini, K.R., Collins J.E., Schultz R.D., Whitlock, R.H., 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol., 12, 685-692.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995, Bayesian data analysis, Chapman & Hall /CRC, Boca Raton, Florida, USA.

Hui, S.L., Walter S.D., 1980. Estimating the error rates of diagnostic tests. Biometrics 36, 167-171.

Kennedy, D.J., Benedictus, G., 2001. Control of *Mycobacterium avium* subsp. *paratuberculosis* infection in agricultural species. Rev. sci. tech. Off. int. Epiz. 20, 151-179.

McKenna, S.L.B., Keefe, G.P. Barkema, H.W., Sockett, D.C., 2005. Evaluation of three ELISAs for *Mycobacterium avium* subsp. *paratuberculosis* using tissue and fecal culture as comparison standards. Vet. Microbiol., 110, 105-111.

Nielsen, S.S., 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subspecies *paratuberculosis* in dairy cattle. J. Vet. Med. B 49, 384-387.

Nielsen SS, Enevoldsen C, Gröhn YT, 2002a. The *Mycobacterium avium* subsp. *paratuberculosis* ELISA response by parity and stage of lactation. Prev. Vet. Med. 54, 1-10.

Nielsen, S.S., Grønbæk, C., Agger, J.F., Houe, H., 2002b. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53, 191-204.

Nielsen, S.S., Gröhn, Y.T., Enevoldsen, C., 2002. Variation of the milk antibody response to paratuberculosis in naturally infected dairy cows. J. Dairy Sci. 85, 2795-2802.

Spiegelhalter, D.J., Thomas, A., Best, N.G., Gilks, W.R., 1995. BUGS: Bayesian inference Using Gibbs Sampling, Version 0.50. MRC Biostatistics Unit, Cambridge, United Kingdom.

Spiegelhalter, D.J., Thomas, A., Best, N.G., 2004, WinBUGS Version 1.4 User Manual, MRC Biostatistics Unit, Cambridge, United Kingdom.

Toft, N., Nielsen, S.S., Jorgensen, E., 2005a. Continuous-data diagnostic tests for paratuberculosis as a multi-stage disease. J. Dairy Sci. 88, 3923-3931.

Toft, N., Jørgensen, E., Højsgaard, S., 2005b. Diagnosing diagnostic tests: evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. Prev. Vet. Med. 68, 19-33.

Toft, N., Innocent, G.T., Gettinby, G., Reid, S.W.J., 2007. Assessing the convergence of Markov Chain Monte Carlo methods: An example from evaluation of diagnostic tests in absence of a gold standard, Prev. Vet. Med. 79, 244-256.

Whitlock, R.H., Rosenberger, A.E., Sweeney, R.W., Spencer, P.A., 1996. Distribution of *M. paratuberculosis* in tissues of cattle from herds infected with Johne's disease. *In*: Chiodini RJ, Hines II ME, Collins MT (eds.): Proceeding of the Fifth International Colloquium on Paratuberculosis, Sept. 29-Oct. 4, 19996, Madison, Wisconsin, USA, p. 168-174.

Paper X

Paratuberculosis in dairy cattle: Variation of the antibody-response in offspring attributable to the dam

Nielsen SS, Gröhn YT, Quaas RL, Agger JF

Published in
Journal of Dairy Science 2002, 85, 406–412

Paratuberculosis in dairy cattle: Variation of the antibody-response in offspring attributable to the dam

S. S. Nielsen[*], Y. T. Gröhn[†], R. L. Quaas[‡], and J. F. Agger[*]

[*] Department of Animal Science and Animal Health, The Royal Veterinary and Agricultural University, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark
[†] Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine and,
[‡] Department of Animal Science, Cornell University, New York 14853, USA

ABSTRACT

The objective of this study was to examine transmission of paratuberculosis in dairy cattle attributable to the dam. Milk samples were collected from 8131 cows in 110 Danish dairy herds. The level of antibodies to *Mycobacterium avium* subspecies *paratuberculosis* was determined by use of an ELISA. Information on dam and sire was obtained from the Danish Cattle database. The following two data sets were analyzed: Data set A contained all cows ≤ 400 days in milk (n = 7410); data set B contained 1056 dam-daughter pairs present simultaneously in herds at the day of sampling. Cows > 400 days in milk were excluded. Linear mixed models were used to obtain variance components for the effect of sire in data set A and the effect of sire and damdaughter pairs in data set B. Models for both data sets A and B included information previously shown to confound antibody level and information of the relative prevalence of paratuberculosis in the herd. In data set A, the effect explained by sire was 1.9%, whereas it was 6.3% in data set B. The effect from dam-daughter pairs was 7.7%. Those effects were all significant. It was concluded that the parental contribution was significant, and both heritability of susceptibility and vertical transmission should be considered in any control programs on paratuberculosis in dairy cattle.

INTRODUCTION

Paratuberculosis is a chronic infectious disease of ruminants caused by *Mycobacterium avium* subsp. *paratuberculosis* (Chiodini et al., 1984). In Denmark (Nielsen et al., 2000) as well as in other parts of the world (Gasteiner et al., 1999; Muskens et al., 2000; NAHMS, 1997) the disease is fairly prevalent. Because the diagnostic tools available suffer from lack of accuracy (e.g., Manning and Collins, 2001), control of the infection relies partly on a thorough understanding of the routes of transmission. Therefore, most recommendations for control of the infection are based on prevention of transmission from other animals in a herd and from a contaminated environment (Collins, 1994; Rossiter and Burhans, 1996; Whitlock et al., 1986). However, transmission can also take place *in utero*, especially in cows manifesting clinical disease or cows that are shedding a high number of bacteria (Seitz et al., 1989; Sweeney et al., 1992a). In addition to *in utero* transmission, a dam is also capable of transmitting the bacteria through close direct contact with the newborn and through feeding milk and colostrum contaminated with feces. Also, infected cows can shed the mycobacteria in milk, not only in the clinical stage (Taylor et al., 1981) but also as asymptomatic cows (Sweeney et al., 1992b), and apparently in higher concentration in colostrum than in milk (Streeter et al., 1995). Vertical transmission attributable to the sire has not been described.

Such transmission is theoretically possible because the mycobacteria have been isolated from the semen and reproductive organs of bulls (Larsen et al., 1981) and from cryopreserved semen intended for AI (Jorge et al., 1998).

Not all infected cows shed the mycobacteria continuously and even fewer show clinical signs (Whitlock and Buergelt, 1996). Identification of some of these animals may rely on detection of an immune response, e.g., detection of antibodies. Antibodies are not protective because *M. avium* subsp. *paratuberculosis* is an intracellular pathogen. Whether both serologically positive and shedding animals have a higher likelihood of giving birth to an animal that will later become seropositive or become a shedder, compared with animals born from nonpositive animals, is not known. The heritability of susceptibility was recently determined to be 0.06 overall (Koets et al., 2000). This estimate apparently includes the effect from vertical transmission, thus suggesting that vertical transmission is not dominating. However, a different estimate of susceptibility, 0.09, was estimated for vaccinated animals in the same study, thus not completely excluding the role of vertical transmission. The importance of vertical transmission thus still has not been established in nonvaccinated cattle.

The objective of this study was to determine the proportion of transmission of paratuberculosis in dairy cattle attributable to the dam with emphasis on vertical transmission.

MATERIALS AND METHODS

Herds and Animals

In 1998, the Danish Dairy Board initiated an integrated milk quality and cattle health study in the Southern part of Jutland. The study region was defined by four postal zip codes. This region was selected because it was assumed to represent the future Danish dairy farm. In this region, there were initially approximately 240 dairy herds. Of these, 110 farmers allowed their herds to be used for studies on infectious diseases including infections with *Salmonella enterica*, *Escherichia coli* O157, *Streptococcus uberis*, and *M. avium* subsp. *paratuberculosis*. These 110 herds were all used in the present study. The herd size, geometrical average bacterial count of bulk tank milk samples, and the geometrical average bulk tank milk SCC of the 110 herds are given in Table 1. Also given in Table 1 are herd size, bacterial count, and SCC for the 130 nonparticipating herds in the region and for the country as a whole. The participating herds were significantly ($P < 0.001$) larger, and the bacterial count was significantly ($P < 0.001$) lower in these herds relative to the herds from the whole country and nonparticipating herds, in the period April 1, 1999, to March 31, 2000. The Mann-Whitney test was used for these nonresponse analyses.

Milk samples were collected from all lactating cows via the routine milk production scheme, from August 1999 to February 2000. One composite sample was collected from each cow. A total of 8131 samples were collected from the 110 herds. In addition, information on parity, days in milk (DIM) age at first calving, milk production at the day of sampling, and the identities of the dam and the sire were collected from the Danish Cattle database. Breeding in all of these herds was through AI. From these observations, the following two data sets were created: Data set A included all observations; data set B included 1106 pairs of observations where a dam and her daughter were both present in the herd as cows at the day of the sampling.

Table 1. Comparison of herd size, and geometrical average of bulk tank milk somatic counts (SSC)[1], and bacterial counts (BC)[2] in the participating 110 herds in the Kongeå region[3] relative to the non-participating herds in the Kongeå region and all herds in Denmark.

Geographical region	Herd size	Geo. SSC	Geo. BC
Kongeå region			
Participating herds			
Minimum	11	108,000	4000
25th percentile	64	256,000	5200
Median	77	269,000	6,500
75th percentile	102	308,000	8300
Maximum	256	447,000	40,000
Mean	85	263,000	8000
Non-participating herds			
Minimum	3	81,000	4000
25th percentile	39	221,000	5900
Median	89	265,000	8400
75th percentile	58	308,000	11700
Maximum	154	426,000	64400
Mean	65	262,000	10300
Denmark, total			
Minimum	1	NA[4]	NA[4]
25th percentile	42	NA[4]	NA[4]
Median	60	NA[4]	NA[4]
75th percentile	84	NA[4]	NA[4]
Maximum	408	NA[4]	NA[4]
Mean	67	251,000	9,000

[1] Geo. SSC = Geometrical average of bulk cell count from April 1, 1999 to March 31, 2000

[2] Geo. BC = Geometrical average of the bacterial count of bulk tank milk samples from April 1, 1999 to March 31, 2000

[3] The Kongeå region was defined by 4 postal zip codes in Southern Jutland, Denmark

[4] NA = Not applicable

Diagnostic Methods

All milk samples were preserved with bronopol immediately after collection. Upon arrival at the laboratory, the samples were centrifuged, the fat fraction removed, and the skim milk frozen for later testing. The content of antibodies to *M. avium* subsp. *paratuberculosis* was estimated by use of an ELISA that has previously been described for serum samples (Nielsen et al., 2001). The test was modified for use on milk samples and it proved as accurate as the serum ELISA with an area under receiver operating characteristic curves of 0.80 for both tests when evaluated by using maximum-likelihood estimation of sensitivity and specificity (Nielsen et al., 2002). All animals in each herd were tested in the laboratory on the same day, thus minimizing any potential laboratory variation in the dam-daughter relationship.

The results from the ELISA were recorded as optical density (OD) values for each cow. A log transformation of the OD value was performed because of an observed skewed frequency distribution. The transformed OD was designated ln(OD).

Statistical Analyses

Linear regression of the ln(OD) was performed by using the mixed procedure of SAS (SAS/STAT, 1997). First, descriptive statistics were obtained. Parity, DIM, age at first calving, and the volume of milk produced on the day of the recording were confounding variables necessary to consider in prediction of the ELISA reading for milk antibodies. This confounding effect is a sequel to the chronic nature of the disease. Hence, these were considered in the statistical analyses. All observations from cows > 400 DIM were excluded from the study because the descriptive statistics indicated fairly different and unpredictable patterns of the OD values at the end of lactation. The number of offspring per sire and the number of dam-daughter pairs per herd were as given in Table 2. After data editing, 7410 cows were analyzed in data set A, and 1056 dam-daughter pairs were analyzed in data set B.

Table 2. Number of offspring per sire in each data set (A and B) for cows ≤ 400 days in milk for 110 Danish dairy herds[†].

Number of sires	Number of offspring per sire	
	Data set A	Data set B
1	{17; 18; 19; 31; 32; 33; 35; 39}	{10; 13; 19}
1	{40; 42; 52; 54; 57; 61; 62; 64}	{22; 26; 30}
1	{68; 72; 81; 82; 91; 97; 114}	{49; 58; 75}
1	{117; 136; 142; 164; 167; 200}	{96}
1	{208; 227; 234; 256; 316}	
2	{21; 22; 24; 26}	{25; 42; 46}
3	{20; 23; 34}	{9; 11; 15; 16; 18}
6	{11; 14; 15}	
7	{13}	
8	{12; 16}	
9		{5; 8}
10		{6}
16	{10}	
22	{9}	
25	{7}	{4}
40	{6}	
42	{8}	
53		{3}
62	{5}	
80	{4}	
83	{3}	
125		{2}
161	{2}	
307		{1}
418	{1}	
133		Missing
620	Missing	

[†]) Interpretation: 1 sire contributed 17 offspring; 1 sire contributed 18 offspring etc.; 2 sires contributed 21 offspring etc.; 418 sires contributed 1 offspring etc.

Subsequent to the descriptive statistics, the relative prevalence (see below) in each herd was estimated; it was used as an approximation of horizontal transmission. Use of the

relative prevalence rather than 'true prevalence' was used in the analysis because the sensitivity of the diagnostic tests is highly dependent on the stage of disease and no 'gold standard' exists to determine the true state of disease in any animal. Because no gold standard exists and because the ELISA has a continuous outcome, the ELISA values were initially standardized. This standardization was performed by using a mixed model with cow characteristics included as fixed effects and herd and laboratory factors as random effects. Such standardization was considered because the factors included have previously been shown to affect the background level in the ELISA.

The fixed effects were as follows: parity (P) as three parity groups, 1 = parity 1, 2 = parity 2, and 3 = parity ≥ 3; DIM (time after calving, T) centered around its grand mean and initially considered as linear, quadratic, cubic, fourth, and fifth power terms; age at first calving (A) centered around its grand mean and initially included as linear, quadratic, and cubic terms; and the dilution effect of milk volume (M) on antibody concentration included as the difference between observed and expected milk production on the jth DIM in the ith parity. All possible interaction terms among P, T, and A were included initially.

Clustering can occur with respect to the ELISA plate in which the sample was tested and herd factors such as mycobacterial background flora and clustering of herds within subsets of plates are possible. Therefore, these were combined as an interaction term and analyzed as a random effect with an unstructured variancecovariance type between samples from the l^{th} herd (H) tested in the m^{th} ELISA plate (E).

Thus, a model was fitted by using Akaike's information criterion (AIC), Schwarz Bayesian criterion, and the likelihood ratio test with a P-value of 0.05. The

resulting model (equation 1) was:

$$y_{ijklm} = u + \mathrm{P}_{ilm} + b_1(\mathrm{T}_{jlm} - \overline{\mathrm{T}}_{.lm}) + b_2(\mathrm{T}_{jlm} - \overline{\mathrm{T}}_{.lm})^2 + b_3(\mathrm{A}_{klm} - \overline{\mathrm{A}}_{.lm}) + b_4\mathrm{M}_{ijklm} + U_{000lm} + \mathrm{H}_l * \mathrm{E}_m + e_{ijklm} \quad (1)$$

where

y_{ijklm} = the logarithm of the OD, ln(OD);

u = the grand mean of ELISA-values;

P_{ilm} = the fixed effect of the i^{th} parity group clustered in the lm^{th} random effect group;

$(\mathrm{T}_{jlm} - \overline{\mathrm{T}}_{.lm})$ = the fixed effect of the j^{th} day in milk centered around the grand mean and clustered in the lm^{th} random effect group;

$(\mathrm{A}_{klm} - \overline{\mathrm{A}}_{.lm})$ = the fixed effect of the k^{th} group of age at 1st calving centered around the grand mean and clustered in the lm^{th} random effect group;

M_{ijklm} = the fixed effect of the observed minus the predicted volume of milk (in kg) on the j^{th} day in the k^{th} parity;

U_{000lm} = the group mean for random effects;

$\mathrm{H}_l*\mathrm{E}_m$ = the combined random effect of the interaction between the l^{th} herd, H, and the m^{th} ELISA-plate, E;

b_1 and b_2 = the parameter estimates for the effect of T;

b_3 = the parameter estimate for the effect of A;

b_4 = the parameter estimate for the effect of M;

e_{ijklm} = the random error;

and where the random effects $\mathrm{H}_l*\mathrm{E}_m$ and e_{jklm} were assumed independent, identically distributed normal with mean 0 and variances σ_{HE}^2 and σ^2, respectively.

After the standardization of the OD readings for each cow, a cutoff was selected to give an overall prevalence of 10%. This was necessary because of the missing gold standard, and

because of this assumption, the prevalences calculated for each herd were designated 'relative prevalence.'

For estimation of the variance components of interest in the study, namely, the contribution of the sire and the dam-daughter relationship, and the contribution of the environment (controlled through the prevalence estimates), models were fitted for the two data sets. The base model was equation 1, despite exclusion of the H*E-term and inclusion of prevalence terms as fixed effects. Thus, the base model here was equation 2, as follows:

$$y_{ijklm} = u + \mathrm{P}_{im} + b_1(\mathrm{T}_{jm} - \overline{\mathrm{T}}_{.m}) + b_2(\mathrm{T}_{jm} - \overline{\mathrm{T}}_{.m})^2 + b_3(\mathrm{A}_{km} - \overline{\mathrm{A}}_{.m}) + b_4\mathrm{M}_{ijm} + b_5\mathrm{C}_l + b_6C_l^2 + b_7C_l^3 + U_{000\,m} + \mathrm{R}_m + e_{ijklm}$$

where y_{ijklm} = ln(OD);

P_{im} = the fixed effect of the i^{th} parity;

$(\mathrm{T}_{jm} - \overline{\mathrm{T}}_{.m})$ = the fixed effect of the j^{th} day in milk centered around the grand mean;

$(\mathrm{A}_{km} - \overline{\mathrm{A}}_{.m})$ = the fixed effect of the k_{th} group of age at 1st calving centered around the grand mean;

M_{ijm} = the fixed effect of the observed minus the predicted volume of milk (kg) on the j^{th} day in the k^{th} parity;

C_l = the fixed effect of contamination of the environment, i.e. the relative prevalence estimated as described above, in the l^{th} herd;

U_{000m} = the group mean for the random effects;

R_m = the random effects in the possible combinations of S+H*E, DD+H*E, S+DD+H*E (for data set B), and H*E, S and S+H*E (for data set A), where DD is the effect of dam-daughter relationship, S is the effect from sire, and H*E is the effect from herd combined with ELISA-plate as in Equation 1;

b_1, b_2, b_3, b_4 and b_5 are the regression coefficients for the terms specified;

e_{ijklm} = the random error;

and where the random effects R_m and e_{jklm} were assumed independent, identically distributed normal with mean 0 and variance components σ_{DD}^2, σ_S^2, σ_{DDS}^2, σ_{HE}^2,σ_{HES}^2 and σ^2.

Evaluation of this model and other models was based on the likelihood ratio test, AIC, and Schwarz Bayesian criterion. The underlying principle of these is that the model with the log likelihood, the Schwarz Bayesian criterion, or the AIC closest to zero is the best model (e.g., theory and use of AIC, reviewed by Bozdogan, 1987). Small differences exist in that Schwarz Bayesian criterion penalizes more on the number of factors included in the model. The variance components of the random effects were estimated by using the restricted maximum likelihood method. Further evaluation of the model fit was done by randomized pairing of cows within herds to provide pseudo-daughter-dam pairs. These random pairs were designated DDr. Plots of the residuals from offspring versus residuals from dams were also inspected to validate the model.

RESULTS

The variance components calculated for various models for data set A are shown in Table 3. The model of interest is the bottom one including the sire term (Table 3, model 2). The other model is shown for comparative purposes only. The variation attributable to sire is 1.9% (95% confidence interval [C.I.]: 1.3–2.5%). For data set B, the variance components for other models are listed in Table 4. Here, model 1 is the "empty" model including only cow characteristics such as parity, DIM and so on, prevalence terms, and correction for clustering because of herd and laboratory procedures. Models 2 and 3a include sire and dam-daughter

pairing, respectively. Model 3b shows the variance component if dams and daughters are paired randomly. This variance is zero as expected. Model 4, with the lower AIC, is the final model. The variance attributable to dam-daughter pairing is 7.7% (95% C.I.: 6.8–15.6%). The variance attributable to sire is 6.3% (95% C.I.: 5.5–12.7%).

Table 3. Variance components from models to predict log ELISA antibody response to *M. avium* subsp. *paratuberculosis*, based on single cow observations (n = 7410) after exclusion of cows > 400 d in milk. Numbers in brackets are standard errors followed by proportion of total variation.

		Variance component				Fit statistics
Model[1]		Herd * ELISA-plate[2]	Sire	Residual	Sum of variation	AIC[3]
1	Cow+Prev+Lab	0.1225*** (0.01548) (22%)		0.4422 (0.00763)	0.5647	15,704
2	Cow+Prev+Lab +Sire	0.1211*** (0.01524) (21%)	0.01069*** (0.003255) (1.9%)	0.4328 (0.00767)	0.5646	15,199

[1] 'Cow' = Cow characteristics (P + T + T^2 + A + M); P = Parity; T = Days in milk centered around the grand mean; A = Age at first calving centered around the grand mean; M = observed minus expected milk yield on the day of sampling); Prev = prevalence within herd included as a linear, a quadratic and a cubic term in the model; Lab = effect of clustering due to laboratory and herd effects (H*E (H = Herd; E = ELISA-plate number)); Cow characteristics and prevalence were included as fixed effects. Other effects were included as random effects. For further details see Equation 2.[2] Herd*ELISA-plate = The interaction term between the effects of herd identity and ELISA-plate identity; [3] AIC = Akaike's Information Criterion. *** $P \leq 0.001$

In addition, it was tested whether the antibody level differed in offspring from dams with different status (alive, slaughtered, or dead in the herd), by inclusion of dam status in equation 1. However, those categories were not found to be significant (data not shown).

DISCUSSION

Establishment of the relative importance of various components of transmission for paratuberculosis is necessary in control measures of the disease. Considering the dam-daughter relationship, the reports of *in utero* transmission (Seitz et al., 1989; Sweeney et al., 1992a) and the reports on the role of *M. avium* subsp. *paratuberculosis* in milk and colostrum (Streeter et al., 1995; Taylor et al., 1981) are worthy of attention along with concerns about direct postnatal contact with infected dams or consumption of fecal-contaminated milk or colostrum. Koets et al. (2000) concluded that vertical transmission was not important whereas a high level of susceptibility was. However, their findings in vaccinated versus nonvaccinated animals are rather contradictory.

In this study, the variation in antibody response of a cow explainable by her dam was estimated to be 7.7%. The effect from sire was 6.3% (Table 4). Both numbers suggest fairly high parental contribution. However, in the large data set (A), the sire effect was only 1.9%, which was significantly different from the 6.3% (95% C.I.: 5.5–12.7%) found in the smaller data set (B). One possible explanation for this difference could be attributed to the "healthy worker survivor effect" (reviewed by Arrighi and Herzt-Picciotto, 1994) in that dams with clinical paratuberculosis (and possibly also subclinical paratuberculosis) are more likely to have been culled, so mostly "healthy" dams remain in the herds. Thus, the daughters of healthy cows should, in general, be more likely to develop antibodies to *M. avium* subsp.

paratuberculosis than those that are daughters of culled dams. However, testing whether cows with high antibody levels were more likely to have died in the herd or have been slaughtered was not significant. Thus, the healthy worker survivor effect was not found in this study.

Table 4. Variance components from models to predict log ELISA antibody response to *M. avium* subsp. *paratuberculosis*, based on dam-offspring pairs (n = 1056) after exclusion of observations > 400 d in milk. In each cell are the covariance parameter estimate, the standard error and the proportion of the total variation.

Model no.	Model[1]	Herd * ELISA-plate[2]	Pair[3]	Sire	Residual	Sum of variation	AIC[4]
1	Cow + Prev+Lab	0.09592***			0.4374	0.5333	4,153
		(0.01595)			(0.01537)		
		18%					
2	Cow + Prev+Lab	0.09404***		0.03373**	0.4108	0.5386	4,140
	+Sire	(0.01580)		(0.01334)	(0.01636)		
		17%		6.3%			
3a	Cow + Prev+Lab	0.1100***	0.04829**		0.3922	0.5505	4,427
	+Dam-daughter pair	(0.01708)	(0.01591)		(0.01928)		
		20%	8.8%				
3b	Cow + Prev+Lab	0.09593***	0		0.4374	0.5333	4,153
	+random pairing of	(0.01595)	0		(0.01537)		
	dam-daughter						
4	Cow + Prev+Lab	0.09254***	0.04151**	0.03374**	0.3705	0.5383	4,136
	+Sire + Dam-daughter	(0.01590)	(0.01629)	(0.01341)	(0.02097)		
	pair	17%	7.7%	6.3%			

[1] 'Cow' = Cow characteristics (P + T + T^2 + A + M); (P = Parity; T = Days in milk centered around the grand mean; A = Age at first calving centered around the grand mean; M = observed minus expected milk yield on the day of sampling); Prev = prevalence within herd included as a linear, a quadratic and a cubic term in the model; Lab = effect of clustering due to laboratory and herd effects (H*E (H = Herd; E = ELISA-plate number)); The effects cow characteristics and prevalence were included as fixed effects. Other effects were included as random effects. For further details see Equation 2. [2] Herd*ELISA-plate = The interaction term between the effects of herd identity and ELISA-plate identity. [3] Pair = Information on antibodies of both dam and offspring. [4] AIC = Akaike's Information Criterion. ** $P \leq 0.01$; *** $P \leq 0.001$

Even with these differences, both models (Tables 3 and 4) suggest an important parental contribution. Whether this is because of vertical transmission or heritability of susceptibility cannot be determined based on the present material. However, in either case, the contribution has a magnitude that deserves attention in a control program. Also, by inclusion of estimates of prevalence, an approximation of the horizontal transmission was attempted. Given that a newborn calf is equally susceptible to infection from the environmental contamination provided by a random cow in a herd as to the contamination of its dam, the estimates seen here account for vertical transmission. This assumption probably does not hold, in general, because close contact to one's own dam is expected.

One drawback of the present study is the use of the immune response as an indicator of infection. Agentdetecting methods could have been used such as was done in the study by Koets et al. (2000) with the underlying assumption that infection causes disease. That is, Koets et al. (2000) used a diagnostic test (Choi, 1997). However, because of a lack of true gold standards for infection with *M. avium* subsp. *paratuberculosis*, the use of a predictive

test such as an immunological status would seem more appropriate in the present context (Choi, 1997).

Still, some cows shedding the bacteria are not detected by using immunological methods. One possible explanation could be passive ingestion and excretion of the mycobacteria without infection that could lead to an immune response (Sweeney et al., 1992c). A different but apparently uninvestigated mechanism in paratuberculosis is the production of immunotolerant offspring similar to that seen with bovine virus diarrhea infection (Moennig and Liess, 1995; Schultz, 1973). Such offspring would be expected if infected dams transfer the mycobacteria in the first and second trimester of their pregnancy. Concurrent agent detection and immunological detection in adult dam-daughter pairs would be necessary to resolve such a hypothesis. Even so, the timing of infection would be hard to determine, although an experimental design as used for bovine virus diarrhea could be used (McClurkin et al., 1984).

Another drawback of the present study is the use of herds that, according to Table 1, are different from the nonparticipating herds and the Danish dairy herds in general. Selection bias could thus have been introduced in the study, but there are no indications that the present selection bias would have a biasing effect on transmission.

In conclusion, the present study indicates that dam effects can explain a significant variation of the antibody level in an offspring. It cannot be determined if the contribution was heritability of susceptibility or vertical transmission. However, because the contribution from the sire is also significant with less probable chance of vertical transmission from sires, both heritability and vertical transmission could be considered in any control programs.

ACKNOWLEDGEMENTS

Thanks are due to the Danish Dairy Board and the Kongeå-project for funding the project and for helping with the organisation of herd contact. Thanks are also due to JDS Editor Steve Washburn and two anonymous reviewers who read and commented upon this paper.

REFERENCES

Arrighi, H.M., and I. Hertz-Picciotto, 1994. The evolving concept of the healthy worker survivor effect. Epidemiology 5: 189-196.

Bozdogan, H., 1987. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika 52: 345-370.

Chiodini, R.J., H.J. van Kruiningen, and R.S. Merkal, 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74: 218-262.

Choi, B.C.K., 1997. Causal modeling to estimate sensitivity and specificity of a test when prevalence changes. Epidemiology 8: 80-86.

Collins, M.T., 1994. Clinical approach to control of bovine paratuberculosis. J. Am. Vet. Med. Assoc. 204, 208-210.

Gasteiner, J., H. Wenzl, K. Fuchs, U. Jark, and W. Baumgartner, 1999. Serological cross sectional study of paratuberculosis in cattle in Austria. J. Vet. Med. B 46, 457-66.

Jorge, M.C., M. Catena, J. Cabodevila, and P. Soto, 1998. Micobacterias en semen bovino criopreservado (In Spanish). Veterinaria Argentina 145: 337-340.

Koets, A.P., G. Adugna, L.L.G. Janss, H.J. van Weering, C.H.J. Kalis, G.H. Wentink, V.P.M.G. Rutten, and Y.H. Schukken, 2000. Genetic variation of susceptibility to

Mycobacterium avium subsp. *paratuberculosis* infection in dairy cattle. J. Dairy Sci. 83: 2702-2708.

Larsen, A.B., O.H.V. Stalheim, D.E. Hughes, L.H. Appell, W.D. Richards, and E.M. Himes, 1981. *Mycobacterium paratuberculosis* in the semen and genital organs of a semen-donor bull. J. Am. Vet. Med. Assoc. 179: 169-171.

Manning, E.J.B., and M.T. Collins, 2001. *Mycobacterium avium* subsp. *paratuberculosis*: pathogen, pathogenesis and diagnosis. Rev. sci. tech. Off. Int. Epiz. 20: 133-150.

McClurkin, A.W., E.T. Littledike, R.C. Cutlip, G.H. Frank, M.F. Coria, and S.R. Bolin, 1984. Production of cattle immunotolerant bovine viral diarrhea virus. Can. J. Comp. Med. 48: 156-161.

Moennig, V., and Liess, B., 1995. Pathogenesis of intrauterine infections with bovine viral diarrhea virus. Vet. Clin. North Am. Food Anim. Pract. 11: 477-487.

Muskens, J., H.W. Barkema, E. Russchen, K. van Maanen, Y.H. Schukken, and D. Bakker, 2000. Prevalence and regional distribution of paratuberculosis in dairy herds in the Netherlands. Vet. Microbiol. 77: 253-261.

NAHMS, 1997. Johne's disease on US dairy operations. USDA:APHIS:VS, CEAH, National Animal Health Monitoring System, Fort Collins, CO, USA, No. N245.1097.

Nielsen, S.S., S.M. Thamsborg, H. Houe, and V. Bitsch, 2000. Bulk-tank milk ELISA antibodies for estimating the prevalence of paratuberculosis in Danish dairy herds. Prev. Vet. Med. 44: 1-7 (with Corrigendum in Prev. Vet. Med. 46: 297).

Nielsen, S.S., H. Houe, S.M. Thamsborg, and V. Bitsch, 2001. Comparison of two enzyme-linked immunsorbent assays for serologic diagnosis of paratuberculosis (Johne's disease) in cattle using different subspecies strains of *Mycobacterium avium*. J. Vet. Diagn. Invest. 13: 164-166.

Nielsen, S.S., C. Grønbæk, J.F. Agger, and H. Houe, 2002. Maximum likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53: 191-204.

Rossiter, C.A., and W.S. Burhans, 1996. Farm-specific approach to paratuberculosis (Johne's disease) control. Vet. Clin. North Am. Food Anim. Pract. 12: 383-415.

SAS Institute, 1997. SAS/STAT: Changes and enhancements through Release 6.12. Cary, NC: Author.

Schultz, R.D., 1973. Developmental aspects of the fetal bovine immune response: A review. Cornell Vet. 63: 507-535.

Seitz, S.E., L.E. Heider, W.D. Hueston, S. Bech-Nielsen, D.M. Rings, and L. Spangler, 1989. Bovine fetal infection with *Mycobacterium paratuberculosis*. J. Am. Vet. Med. Assoc. 194: 1423-1426.

Streeter, R.N., G.F. Hoffsis, S. Bech-Nielsen, W.P. Shulaw, and M. Rings, 1995. Isolation of *Mycobacterium paratuberculosis* from colostrum and milk of subclinically infected cows. Am. J. Vet. Res. 56: 1322-1324.

Sweeney, R.W., R.H. Whitlock, and A.E. Rosenberger, 1992a. *Mycobacterium paratuberculosis* isolated from fetuses of infected cows not manifesting signs of the disease. Am. J. Vet. Res. 53: 477-480.

Sweeney, R.W., R.H. Whitlock, and A.E. Rosenberger, 1992b. *Mycobacterium paratuberculosis* cultured from milk and supramammary lymph nodes of infected asymptomatic cows. J. Clin. Microbiol. 30: 166-171.

Sweeney, R.W., R.H. Whitlock, A.N. Hamir, and A.E. Rosenberger, 1992c. Isolation of *Mycobacterium paratuberculosis* after oral inoculation in uninfected cattle. Am. J. Vet. Res. 53: 1312-1314.

Taylor, T.K., C.R. Wilks C. R., and D.S. McQueen, 1981. Isolation of *Mycobacterium paratuberculosis* from the milk of a cow with Johne's disease. Vet. Rec. 109: 532-533.

Whitlock, R.H., and C. Buergelt, 1996. Preclinical and clinical manifestations of paratuberculosis (including pathology). Vet. Clin. North Am. Food Anim. Pract. 12: 345-356.

Whitlock, R.H., L.T. Hutchinson, L.T. Glickman, R. Meinersmann, C. Rossiter, J. Bruce, R. Merkal, and J. Dick, 1986. Paratuberculosis (Johne's disease) update. Bovine Practitioner 21: 24-30.

Paper XI

Colostrum and milk as risk factors for infection with *Mycobacterium avium* subsp. *paratuberculosis* in dairy cattle

Nielsen SS, Bjerre H, Toft N

Published in
Journal of Dairy Science 2008, 91, 4610–4615

Colostrum and milk as risk factors for infection with *Mycobacterium avium* subspecies *paratuberculosis* in dairy cattle

S.S. Nielsen[*1], H. Bjerre[*], N. Toft[*]

[*]Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark

ABSTRACT

Mycobacterium avium ssp. *paratuberculosis* (MAP) infections cause major losses to the dairy industry. Transmission of MAP occurs primarily via feces and *in utero*, but MAP can also be excreted in colostrum and milk. The objective of this study was to determine whether colostrum and milk fed to calves are important risk factors for infection with MAP. A questionnaire was sent to 1,050 farms participating in the Danish control program on paratuberculosis in early 2007. Details on practices regarding colostrum and milk feeding between 1999 and 2006 were obtained from 808 (77%) herds. Nine vaccinated herds were excluded. Information on MAP antibody-ELISA results, date of birth, and herd of birth of 93,994 animals was obtained from the Danish Cattle Database. A 2-level logistic regression model was fitted with a dichotomized ELISA response, with outcome, age, source of colostrum, and milk as fixed effects, and herd as a random effect. Animals fed colostrum from multiple cows had an odds ratio of 1.24 of being ELISA positive compared with animals fed colostrum from their own dam only. Calves suckling with foster cows had an odds ratio of 2.01 of being ELISA positive compared with calves fed milk replacer. Feeding bulk tank milk and pooled milk from cows with high somatic cell counts did not increase the risk of being ELISA positive. Overall, the results of the study suggested that source of milk was not of great importance for the transmission of MAP, but colostrum should be fed only from the dam of that calf.

INTRODUCTION

Infections with *Mycobacterium avium* ssp. *paratuberculosis* (MAP) cause significant losses to the dairy industry, primarily because of reduced milk production and premature culling (Ott et al., 1999). Breaking transmission routes appears to be the most cost-effective option for reducing the prevalence of MAP infection in a herd (Groenendaal et al., 2002; Kudahl et al., 2007). Young calves are considered to be more susceptible to infection than older calves (Taylor, 1953), but adult cows can also be infected (Doyle, 1953). Transmission occurs primarily via transmission *in utero* (Whittington and Windsor, 2009) or via ingestion of MAP. The latter occurs after contamination of the environment of the calves or ingestion of MAP shed in milk (reviewed in Sweeney, 1996).

Multiple management-related risk factors for infection with MAP have been identified, such as housing in bed stalls instead of tie stalls, poor hygiene standards in the feeding area of calving pens, low amounts of straw in the calving area, high animal density among young stock (Nielsen and Toft, 2007), manure buildup in calving areas (Johnson-Ifearulundu and Kaneene, 1998; Berghaus et al., 2005), and being born to a sero-positive dam (Aly and Thurmond, 2005). In addition to excretion in feces, MAP can also be excreted in colostrum and milk (Alexejeff-Goloff, 1929; Taylor et al., 1981; Sweeney et al., 1992; Streeter et al., 1995). However, excretion appears to be correlated with stage of infection (Sweeney et al.,

1992), and only a few infected cows would be expected to shed MAP in milk in the early stages of infection. It is also possible that milk can be contaminated with feces containing MAP or that teats can be contaminated with MAP, and a calf may then ingest MAP at milk feeding without MAP actually being excreted in the milk (Sweeney, 1996). Although milk is a putative risk factor, feeding of milk to susceptible calves may not result in an increased risk of infection.

Calves are fed colostrum during the first days of life to obtain IgG antibodies from their dams. Subsequently, they may either be fed milk replacer, waste milk from cows with high concentrations of SCC, milk containing antibiotics, or bulk tank milk, or they may suckle with foster cows. Milk containing antibiotic residues cannot be used for human consumption, and milk from cows with high SCC are often withheld from the bulk tank milk, because bulk tank milk with >400,000 cells/mL cannot be delivered to dairies in the European Union (European Economic Community, 1992). Therefore, such waste milk is often pooled and used for calves. In 26 Danish dairy herds, cows with antibodies to MAP often had high SCC (Baptista et al., 2008), and antibodies to MAP are an indicator of shedding of MAP in feces (Nielsen and Toft, 2006). If milk from MAP-infected cows is an important risk factor, use of milk with high SCC concentrations may have a major effect on the spread of MAP. In addition, the use of pooled milk or colostrum potentially increases the number of infectious animals contributing MAP to a batch of milk, and therefore the risk of transmitting MAP to more calves. Our objective was to determine, in a population-based study, the effect of milk and colostrum source fed to calves on the future risk of developing antibodies to MAP.

MATERIALS AND METHODS

Study Design and Data Collection

A voluntary control program on paratuberculosis for dairy cattle was established in Denmark in February 2006 (Nielsen, 2007). Participating herds are expected to conduct 4 annual ELISA tests of all lactating animals, and 1,050 (19%) of the 5,387 dairy herds in Denmark were registered in the program as of February 1, 2007.

A one-page questionnaire was sent to the managers of 1,050 herds in March 2007, and a reminder was sent 1 mo later to those that had not responded. In the questionnaire, the herd manager was asked the following questions:

a) Has vaccination against paratuberculosis been conducted in the past 8 yr?
b) State the period for which you have knowledge of the milk-feeding practices in the herd.
c) What source of colostrum was used for heifer calves? Options: 1) from own dam; 2) pooled from several dams; 3) combinations of 1 and 2.
d) What type of milk was used for heifer calves after the colostrum period? Options: 1) milk replacer; 2) bulk tank milk; 3) pooled milk from cows with high SCC; 4) bulk tank milk if insufficient milk from cows with high SCC; 5) calves with foster cows.

For each question, the herd manager should indicate the period during which the practice was performed. The semesters were January to June 1999, July to December 1999, January to June 2000, and so on to July to December 2006.

Additional information regarding vaccination was also obtained from the Veterinary Food Administration (Mørkhøj, Denmark), which issues vaccination permissions, and from the National Veterinary Institute (Copenhagen, Denmark), which sells the vaccines.

Vaccinated herds were excluded from the study, because vaccination can result in antibodies, which are not caused by natural MAP infection.

Information regarding MAP antibody ELISA results, date of birth, herd of birth of animals, and geometric mean SCC (Herd gSCC) and geometric mean bacterial count (Herd gBC) of bulk tank milk samples was obtained from the Danish Cattle Database. For each animal, the types of colostrum and milk it had received were based on the date of birth, the herd of birth, and the data from the questionnaire for each herd. Animals born in herds for which questionnaire data did not exist were excluded.

Paratuberculosis Diagnosis

Mycobacterium avium ssp. *paratuberculosis* antibody ELISA results in the Danish Cattle Database are based on milk samples from the Danish milk recording system, tested by using an in-house ELISA (Nielsen, 2002) at Eurofins Steins Laboratory (Holstebro, Denmark). The ELISA readings were recorded as corrected optical densities, and the cutoff used for classifying an animal as antibody positive was corrected optical densities >0.3, as used in the Danish control program. Sensitivity and specificity of the ELISA varies with age because of the chronic nature of the infection (Nielsen and Toft, 2006); therefore, age was included as a covariate in the statistical analyses.

Statistical Methods

Descriptive Statistics. The proportion of responders was calculated as the number of farm managers returning the questionnaire among all participants in the control program on March 1, 2007. The animals belonging to farms responding to the questionnaire constituted the study population. A nonrespondent analysis was conducted describing herd size, Herd gSCC, and Herd gBC for responders, nonresponders, and herds not in the control program. Differences between populations were evaluated by pair-wise comparisons with the Mann-Whitney test by using the PROC NPAR1WAY in SAS version 9.1 (SAS Institute, Cary, NC). In addition to the nonrespondent analysis, the distribution of antibody-positive animals in each of the colostrum and milk groups as well as in 5 age groups was calculated.

Multilevel logistic regression. Multilevel logistic regression was carried out by using the following model:

$$\ln\left(\frac{P(E^{+}_{ijk})}{1-P(E^{+}_{ijk})}\right) = U_0 + \mathrm{C}_i + \mathrm{M}_j + \mathrm{C}_i \times \mathrm{M}_j + \mathrm{A}_k + \nu_{oh} \quad (1)$$

where $P(E^{+}_{ijk})$ = the probability that the dichotomised ELISA-result was positive;

U_0= the baseline probability of testing positive in ELISA;
C_i = the effect of the i^{th} colostrum source (1=colostrum from own dam only; 2=partly colostrum from own dam, partly pooled colostrum; 3=pooled colostrum only);
M_j = the effect of the j^{th} milk source (1 = milk replacer, 2=bulk tank milk; 3=pooled milk from cows with high SCC; 4=bulk tank milk if insufficient milk from cows with high SCC; 5=calves were suckling milk from foster cows);
A_k = the effect of the k^{th} age group (1 = <3 yr; 2 = 3 to 4 yr; 3 = 4 to 5 yr; 4 = 5 to 6 yr; 5 = >6 yr of age); and
v_{0h} = the random intercept for herd *h* with compound structure covariance.
Results were deemed significant if $P < 0.05$, and model terms were excluded if they were nonsignificant. The random-effects multilevel logistic regression model was analyzed by

using PROC GLIMMIX in SAS version 9.1, and odds ratios (OR) and 95% confidence intervals were calculated for comparison of the different groups.

RESULTS

Respondent and non-respondent analysis

Questionnaires were received from 808 (77%) of the 1,050 herds enrolled in the Danish paratuberculosis control program. Of these 808 herds, 9 had been vaccinated during the study period. All animals born during the vaccination periods were excluded. Distributions of Herd gSCC and Herd gBC of the bulk tank milk samples, and herd size of responders, nonresponders, and nonparticipating Danish dairy herds are given in Table 1. No significant differences were seen in Herd gSCC and Herd gBC between responders and nonresponders. The statistical analysis indicated that the herd size for responders (median = 130) was significantly smaller than that for nonresponders (median = 141). However, this difference was probably not biologically important. Compared with herds not participating in the control program, the Herd gSCC and Herd gBC were significantly lower ($P < 0.001$) among responders. Nevertheless, the medians appeared biologically similar (Herd gSCC: 226,571 for responders and 244,153 for nonparticipants; Herd gBC: 5,780 for responders and 6,578 for nonparticipants). The median herd size among responders (130 cows) was almost one-third larger than the median herd size among nonparticipants (99 cows; $P < 0.001$).

Table 1. Comparison of herd size, and geometric mean of bulk tank milk somatic cell counts (gSCC),[1] and bacterial counts (gBC)2 in 808 responding and 242 nonresponding herds in the Danish paratuberculosis control program, and 4,337 herds not in the Danish paratuberculosis control program

Group	Herd size	gSSC	gBC
In control program			
Responders			
Minimum	15	87,017	3,101
25th percentile	95	183,854	4,492
Median	130	226,571	5,780
75th percentile	165	274,726	8,428
Maximum	524	390,773	36,617
Non-responders			
Minimum	12	78,708	3,154
25th percentile	110	193,755	4,588
Median	141	222,886	5,811
75th percentile	187	276,106	9,190
Maximum	1304	390,645	36,201
Not in control program			
Minimum	1	42,580	3,000
25th percentile	56	195,149	4,861
Median	99	244,153	6,578
75th percentile	144	296,383	10,355
Maximum	901	1,019,000	435,806

[1] gSSC = Geometric mean of bulk cell count from June 1, 2006 to May 31, 2007; [2] gBC = Geometric mean of the bacterial count of bulk tank milk samples from June 1, 2006 to May 31, 2007

Descriptive results

Information on MAP antibody ELISA and sources of colostrum and milk were available for 93,994 cows. Cross-tabulations of ELISA results with source of colostrum, source of milk, and age at testing are given in Table 2.

Table 2. Cross-tabulations of number of observations of paratuberculosis ELISA results for cows fed with colostrum and milk from different sources when they were calves

	Milk ELISA		Total
	+	-	
Colostrum source			
Own dam	9,110	54,198	63,308
Multiple cows	2,252	12,244	14,496
Combination of own dam and multiple cows	2,266	13,924	16,190
Milk source			
Milk replacer	1,833	11,403	13,236
Bulk tank milk only	894	5,337	6,231
Pooled milk from high SCC cows	4,298	24,587	28,885
Bulk tank milk if no milk from cows with high SCC	4,176	24,451	28,627
Milk replacer and pooled milk from high SCC cows	2,247	13,834	16,081
Suckling with foster cows	180	754	934
Age group			
< 3 years	1,676	27,053	28,729
3-4 years	4,006	23,411	27,417
4-5 years	3,883	14,585	18,468
5-6 years	2,249	8,367	10,616
>6 years	1,814	6,950	8,764

Multilevel logistic regression

Pearson's chi-square fit statistics suggested a reasonable fit of the model and results of multilevel logistic regression are given in Table 3. Interaction between source of milk and source of colostrum was not significant. Age group ($P < 0.001$), source of milk ($P = 0.012$), and source of colostrum ($P = 0.005$) were significant overall. However, compared with milk replacer, only feeding milk from a combination of cows with high SCC and bulk tank milk, compared with feeding milk replacer and housing calves with foster cows, gave significantly greater OR. Pooling colostrum gave a greater OR of ELISA positivity than using colostrum from the dam of the calf. All OR were low (1.25), except for calves left with foster cows (OR = 2.0 compared with feeding milk replacer; Table 3).

Table 3. Multilevel logistic regression comparing odds of testing positive in antibody ELISA to *Mycobacterium avium* ssp. *paratuberculosis* in cows fed colostrum and milk from different sources

Stratum	Regression coefficient	Standard error	OR (95% CI) [1]	P-value [2]
Intercept	-2.8996	0.0605		<0.0001
Source of colostrum				0.0050
From own dam only (reference)	0	-	1.000	-
From multiple cows	0.2174	0.0673	1.243 (1.089; 1.418)	0.0012
From own dam and multiple other cows	0.0027	0.0653	1.003 (0.882; 1.140)	0.9669
Milk source				0.0120
Milk replacer (reference)	0	-	1.000	-
Bulk tank milk only	0.0676	0.0909	1.070 (0.895; 1.279)	0.4575
Pooled milk from high SCC cows	0.0928	0.0676	1.097 (0.961; 1.153)	0.1701
Bulk tank milk if no milk from cows with high SCC	0.1280	0.0653	1.137 (1.000; 1.292)	0.0498
Milk replacer and pooled milk from high SCC cows	0.0542	0.0757	1.056 (0.910; 1.225)	0.4740
Suckling with foster cows	0.6996	0.1962	2.012 (1.370; 2.956)	0.0004
Age group				<0.0001
< 3 years (reference)	0	-	1.000	-
3-4 years	1.1116	0.0285	3.039 (2.874; 3.214)	<0.0001
4-5 years	1.7269	0.0292	5.623 (5.311; 5.954)	<0.0001
5-6 years	1.8070	0.0326	6.092 (5.715; 6.494)	<0.0001
>6 years	1.8533	0.0342	6.381 (5.967; 6.823)	<0.0001
Herd	0.4020	0.0288	-	<0.0001

[1]) OR=odds ratio; CI= confidence interval

[2]) P-value: The probability that the odds of being ELISA-positive is different from the reference group

DISCUSSION

Feeding colostrum and milk from MAP-infected cows has been suspected as being a risk factor for MAP transmission (Sweeney, 1996), particularly if milk is used from cows in the final stages of infection, because MAP can be excreted in milk (Sweeney et al., 1992). However, the importance of colostrum and milk for feeding of calves in transmission of MAP in the general dairy population has not been described previously. We demonstrated that there is an increased probability of testing positive with an antibody test when colostrum from multiple cows is used, as well as when pooled milk from cows with high SCC counts is combined with feeding of bulk tank milk. A likely explanation is that pooling of milk will increase the probability that colostrum or milk containing infective concentrations of MAP is fed to susceptible calves.

All milk sources yielded OR >1, but were either non-significant or only borderline significant, except when the source of milk was foster cows, which was significant. The results suggest that milk from MAP-infected cows is a risk factor, but the impact of milk feeding is minor. Had milk been an important risk factor, a larger effect on the OR would have been reflected in our large data set. The finding that animals that had suckled a foster cow during calfhood had a very high risk of testing ELISA positive compared with calves fed milk replacer indirectly supports the conclusion that milk was not an important risk factor. Calves kept with a MAP-infected foster cow would be at risk not only from the milk feeding, but also from staying for a long time in the environment of this cow and from suckling on teats contaminated with feces. Therefore, a mixed effect of transmission via milk and feces would be expected in this group of animals. Considering the much greater OR obtained from

staying with only one or a few cows of unknown MAP infection status, compared with a random calf fed milk replacer and subjected to other risk factors, exposure to the environment of the cows is apparently a much greater risk factor than being fed the milk itself. However, milk contaminated with feces containing MAP should also be considered part of this environmental risk.

The major strengths of our study were the large sample size and the retrospective data. A large sample size may increase the risk of finding statistically significant results where biological significance is vague. Therefore, the statistical significance found should be interpreted in conjunction with the size of the OR. Drawbacks include differences in the study population compared with the remaining Danish dairy cattle population as described in the respondent and nonrespondent analysis; use of ELISA, which is not a perfect test for detection of MAP-infected animals; and use of questionnaire data. Lack of representativeness of the study population, as suggested by the nonrespondent analysis, does not cause major concern, particularly because the number of herds and animals included was large. Increasing the sensitivity of ELISA by reducing the cutoff resulted in a slightly increased OR for calves kept with foster cows, but not for the other OR estimates or the significance of the results (data not shown). Misclassification error by the ELISA was therefore deemed to have only a vague effect on the estimates. Recall bias of the farmers could have influenced the results because they were asked to remember how they fed their calves up to 8 yr earlier. Therefore, the questionnaire was kept very short and addressed primarily questions regarding milk and colostrum feeding. However, on many farms, colostrum- and milk-feeding routines are not changed frequently, and recall bias was not considered a major concern. Should there be an effect of recall bias, it is likely that it would effect the size of the standard errors and thereby the lack of significance of some of the nonsignificant groups. The main potential confounder, namely, the risk via feces from cows, especially around calving, was not assessed. Although recall bias regarding milk feeding was considered to be low, we were concerned that we would be unable to obtain precise information about management regarding calving, because management practices often differ from one calving to another, and often also over years. Therefore, the estimates presented may be confounded, but estimation of the potential confounding was not possible.

We did not include the option "Use of colostrum from only one cow other than the dam" in the questionnaire. During the study period, this practice was not common; however, since the beginning of the Danish control program, it has become more so, for example, as part of the use of colostrum banks. If colostrum from banks is used for colostrum feeding, it would be considered important if the colostrum is from one or more cows, because the risk that a batch of colostrum contains MAP will increase with the number of contributors. Therefore, the results from this study would be applicable also to situations in which colostrum from one cow to one calf is used, and not only from the dam to the calf of that dam.

To summarize, pooling of milk and pooling of colostrum appear be risk factors for transmission of MAP, but their effects are small. Milk and colostrum from cows shedding high bacterial counts or in the clinical stage of infection should still not be used for calf feeding, but it may seem reasonable to relax management procedures aimed at reducing transmission via milk if the efforts are instead directed toward avoiding transmission around calving and in situations in which calves are being exposed to manure from cows. We strongly recommend that colostrum is fed only from a dam to her own calf or calves, or from one cow to one calf, and that pooling of colostrum be avoided.

CONCLUSION

Results from this study demonstrate that the source of milk and colostrum can be risk factors in the development of antibodies to MAP, which are indicators of MAP infection; however, they probably play only a minor role in the transmission within the dairy herds. Colostrum from the dam of the calf resulted in a reduced risk of MAP infection compared with colostrum from multiple cows. Milk from other sources, excluding milk replacer, resulted in slightly greater odds of being MAP infected than did feeding of milk replacer, but the effects were generally small. Calves kept with foster cows had the greatest risk of becoming MAP infected.

ACKNOWLEDGEMENTS

This study was supported by the Danish Directorate of Food, Fisheries and Agro Business (Copenhagen, Denmark; DFFE, J.nr. 3412-06-01550), the Danish Milk Levy Fund (c/o Danish Dairy Board, Århus, Denmark), and the Danish Cattle Levy Fund (c/o Danish Meat Association, Copenhagen, Denmark).

REFERENCES

Alexejeff-Goloff, N. A. 1929. Zur Frage der Pathogenese und Bazillenausscheidung bei Rinderparatuberkulose. Zeitschrift für Infektionskrankheiten, parasitäre Krankheiten und Hygiene der Haustiere 36: 313-317.

Aly, S. S., and M. C. Thurmond. 2005. Evaluation of *Mycobacterium avium* subsp. *paratuberculosis* infection of dairy cows attributable to infection status of the dam. J. Am. Vet. Med. Assoc. 227:450-454.

Anon. 1992. Council Directive 92/46/EEC of 16 June 1992 laying down the health rules for the production and placing on the market of raw milk, heat-treated milk and milk-based products. Off. J. L 268, 14/09/1992: 1–31.

Baptista, F. M., S. S. Nielsen, and N. Toft. 2008. Association between the presence of antibodies to *Mycobacterium avium* subspecies *paratuberculosis* and somatic cell count. J. Dairy Sci. 91: 109-18.

Berghaus, R. D., J .E. Lombard, I. A. Gardner, and T. B. Farver. 2005. Factor analysis of a Johne's disease risk assessment questionnaire with evaluation of factor scores and a subset of original questions as predictors of observed clinical paratuberculosis. Prev. Vet. Med. 72: 291-309.

Doyle, T. M. 1953. Susceptibility to Johne's disease in relation to age. Vet. Rec. 65:363–365.

Groenendaal, H., M. Nielen, A. W. Jalvingh, S. H. Horst, D. T. Galligan, and J. W. Hesselink. 2002. A simulation of Johne's disease control. Prev. Vet. Med. 54:225-245.

Johnson-Ifearulundu, Y. J., and J. B. Kaneene. 1998. Management-related risk factors for *M. paratuberculosis* infection in Michigan, USA, dairy herds. Prev Vet Med. 37: 41-54.

Kudahl, A. B., S. Østergaard, J. T. Sørensen, and S. S. Nielsen. 2007. A stochastic model simulating paratuberculosis in a dairy herd. Prev. Vet. Med. 78: 97-117.

Nielsen, S. S. 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subspecies *paratuberculosis* in dairy cattle. J. Vet. Med. B 49:384-387.

Nielsen, S. S. 2007. Danish control programme for bovine paratuberculosis. Cattle Practice, 15: 161-168.

Nielsen, S. S., and Toft, N. 2006. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. J. Dairy Sci. 89: 569-579.

Nielsen, S. S., and Toft, N. 2007. Assessment of management-related risk factors for paratuberculosis in Danish dairy herds using Bayesian mixture models. Prev. Vet. Med. 81: 306-317.

Ott, S. L., S. J. Wells, and B. A. Wagner. 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40: 179-192.

Streeter, R. N., G. F. Hoffsis, S. Bech-Nielsen, W. P. Shulaw, and D. M. Rings. 1995. Isolation of *Mycobacterium paratuberculosis* from colostrum and milk of subclinically infected cows. Am. J. Vet. Res. 56: 1322-4.

Sweeney, R. W. 1996. Transmission of paratuberculosis. Vet. Clin. North Am.Food Anim. Pract. 12: 305-312.

Sweeney, R. W., R. H. Whitlock, and A. E. Rosenberger. 1992. *Mycobacterium paratuberculosis* cultured from milk and supramammary lymph nodes of infected asymptomatic cows. J. Clin. Microbiol. 30: 166-171.

Taylor, A. W. 1953. Experimental Johne’s disease in cattle. J. Comp. Pathol. 63:355–367.

Taylor, T. K., C. R. Wilks, and D. S. McQueen. 1981. Isolation of *Mycobacterium paratuberculosis* from the milk of a cow with Johne's disease. Vet. Rec. 109: 532-533.

Whittington, R. J., and P. A. Windsor. 2009. *In utero* infection of cattle with *Mycobacterium avium* subsp. *paratuberculosis*: A critical review and meta-analysis. Vet. J. 179: 60-69.

Paper XII

Assessment of management-related risk factors for paratuberculosis in Danish dairy herds using Bayesian mixture models

Nielsen SS, Toft N

Published in
Preventive Veterinary Medicine 2007, 81, 306–317

Assessment of management-related risk factors for paratuberculosis in Danish dairy herds using Bayesian mixture models

S. S. Nielsen, N. Toft

Department of Large Animal Sciences, Faculty of Life Science, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark

ABSTRACT

Transmission of *Mycobacterium avium* subsp. *paratuberculosis* (Map) to susceptible animals is primarily considered to occur via faeces and milk originating from infectious cows. However, studies of factors resulting in increased transmission of Map are difficult to perform due to a long and unpredictable incubation period and inaccurate diagnostic tests. A multi-level Bayesian mixture model has been shown to predict the infection status of an individual cow more precisely than traditional cut-off based methods used for interpretation of diagnostic test-information, thereby increasing the precision of the diagnostic information.

The objective of our study was to assess management-related risk factors for within-herd transmission of Map. Management-related risk factors were recorded in 97 Danish dairy herds. Twenty-six months following that recording, the antibody status of all lactating cows (n = 7410) in the same herds was measured by the use of an ELISA. A multi-level Bayesian mixture model was used to assess the association between the probability of infection of individual cows and 41 herd-level management-related risk factors using univariable analyses. In this model, the continuous OD value was used to estimate the probability of infection, corrected for known animal covariates and laboratory factors. The statistical significance of the potential risk factors was assessed by calculating odds ratios and their 95% credibility posterior intervals.

Four significant risk factors were identified: housing of cows in bed stalls compared to housing in tie stalls; low level of hygiene in the feeding area of calving areas; low amounts of straw in the bedding of the calving area; high animal density among young stock >12 months of age. Surprisingly, the hygiene level in the calving area was not found to affect the odds of infection.

INTRODUCTION

Paratuberculosis is a chronic, slowly developing infection caused by *Mycobacterium avium* subsp. *paratuberculosis* (Map) primarily affecting ruminants. Infection is believed to mainly occur in calfhood, with age-resistance occurring at 3–6 months (Taylor, 1953). According to simulation studies, test-and-cull strategies alone do not appear to be efficient (Groenendaal et al., 2002); therefore, control of the infection should focus on interrupting transmission routes. Primary transmission routes are via faecal contamination of calves' environment by infectious cows, and the shedding of Map in milk by infectious cows, which is subsequently fed to susceptible calves. Adult cattle are the primary sources of Map, with the highest proportion of animals shedding bacteria at 3–5 years of age (Kalis et al., 1999). Housing of infectious calves with susceptible calves has resulted in infection of the latter in experimental studies (van Roermund et al., 2007). It has been established that shedding of Map in milk occurs (Sweeney et al., 1992a), however, the relative contribution of infection via contaminated milk versus contaminated faeces has not yet been established. Other routes of

transmission, i.e. in uteri (Sweeney et al., 1992b) and via contaminated semen may also be possible, but these transmission mechanisms probably play a minor role (Anon., 2004; Mortensen et al., 2004). Therefore, transmission of Map may be greatly reduced if milk and faeces from infectious cows are kept away from susceptible animals, i.e. younger calves. This implies that important transmission routes could be broken by proper implementation of management procedures aimed at protecting calves from contact with infectious cows. However, little scientific evidence from observational studies exists to support this hypothesis. Presently, the approach to assess management-related risk factors for paratuberculosis has been to use the herd infection status as the outcome (established by serological testing of a sample of animals: one or more positive animals implied that the herd was positive), thereby ignoring the differences in within-herd prevalence (Wells and Wagner, 2000; van Weering et al., 2005). Except for purchase of animals and herd size, few other risk factors associated with the presence of sero-positive animals were found. Both studies included a time lag between diagnostic sampling and recording of the investigated management factors, thereby partly managing the slow development of Map-infection.

Ignoring variation in within-herd prevalence implies that this variation should be attributed to random fluctuations. However, it appears as if some infected farms manage the infection better than others (Nielsen et al., 2007). As a consequence, some farms have a consistently high prevalence across time, while others have a consistently low prevalence. This could be hypothesised that low prevalence farms are able to better control within-herd transmission. To explore this further, it is necessary to classify animals rather than herds as infected or not-infected. However, available diagnostic tests for paratuberculosis at the animal level are notoriously poor with low sensitivity, thereby resulting in an unacceptable degree of misclassification of individual animals if a traditional approach to interpretation of test results is used. The outcome of a serological test such as an ELISA is a continuous optical density (OD). For paratuberculosis, animal covariates such as age, as well as laboratory factors, are known to influence the OD as well as infection status. Further, it has been shown that the higher the OD, the higher the probability of shedding, i.e. progression of the infection (Collins, 2002; Toft et al., 2005). Therefore, using the continuous OD rather than a dichotomised value, while accounting for covariates, could improve the diagnostic power of an ELISA and allow estimation of the probability of infection at the animal level.

In Nielsen et al. (2007), it was demonstrated that known covariates such as variation caused by laboratory factors and age of the test subject could be included in a multi-level Bayesian mixture model using the continuous OD of a milk-based ELISA to estimate the probability of infection of individual cows. This mixture model was shown to be more precise in establishing the within-herd prevalence than the traditionally used cut-off based model. Furthermore, the model does not classify animals as infected/not-infected per se, thereby circumventing the problem with misclassification. Adapting this mixture model to analyse the effect of assumed management-related risk factors on the probability of infection at the animal level would therefore be a possible approach to more quantitatively assess the impact of management-related risk factors, thus providing a means of analysing differences between herds in a high risk area where purchase can be hypothesised to maintain/reintroduce the infection.

The objective of the present study was to assess the impact of management-related herd-level risk factors on the prevalence of paratuberculosis infection in 97 Danish dairy

herds, using a multi-level Bayesian mixture model of the individual probability of infection for all lactating cows in the herds.

MATERIALS AND METHODS

Herds, study period and test-scheme

The herds studied were a part of an integrated herd-health and milk quality project initiated by the Danish Dairy Board (Andersen et al., 2000). Among the approximately 240 dairy herds present in the research region, defined by 4 adjacent postal areas in Southern Jutland, Denmark, 110 farmers allowed for their herd to be used for studies on paratuberculosis, mastitis, *E. coli* 0157 and *Salmonella enterica* subsp. *enterica*. Ten herds withdrew from the study during the study period, 8 because they ceased production. Investigations relating to paratuberculosis were conducted in 100 herds, but management factors were not recorded in 3 of these, so the sample used for this study consisted of 97 herds. There was no official paratuberculosis control programme operating in the study period. A non-correspondence analysis showed that the participating herds were different in herd size (median: 77 cow-years; mean: 85 cow-years) than the non-participating herds (median: 89 cow-years; mean: 65 cow-years) (Nielsen et al., 2002).

Recording of herd and management-related factors

Herd information and management-related factors, which were considered as potential risk factors for transmission of Map, were recorded once in each herd from August 1999 to December 1999. Four milk quality advisors from the Danish Dairy Board were trained together to record various factors related to milk quality, housing and other management factors. The training was performed to improve inter-observer agreement, but there was no formal assessment of their agreement. The type of housing for each of 6 age groups, and a total of 35 management-related factors were recorded. The recordings were done using either a nominal (e.g., type of stable), a dichotomous (e.g., cows on pasture or not) or an ordinal scale with scores from 1 to 5, with 1 being the best and 5 the worst, for factors relating to hygiene, density and handling. The observations were made for different age-groups/risk areas: (a) facilities and management at calving and calving area; (b) facilities and management of calves <2 months of age; (c) facilities and management of calves 2–12 months of age; (d) facilities and management of young stock >12 months of age (prior to 1st calving); (e) facilities and management of cows (after 1st calving). The recorded management factors are listed in Table 1a and Table 1b.

Evaluation of infection status

Approximately 26 months after the recording of the management factors, all lactating cows were tested with a milk ELISA for detection of antibodies to Map. The milk samples were obtained from the routine milk-recording scheme. The milk-ELISA used was an in-house test based on a commercially available antigen (Allied Monitor, Fayette, Missouri, USA), with inclusion of a pre-absorption step using *M. phlei* (Nielsen, 2002). When used as a traditionally dichotomised test, it has been evaluated to have a sensitivity of detection of infection of 0.06 for cows 2 years of age, increasing to 0.50 for cows 5 years of age (Nielsen and Toft, 2006). The specificity was established to 0.997 for animals 2 years of age and 0.93 for animal that were 5 years of age. These estimates were obtained in infected herds, thereby explaining the low specificity assuming that transient infections occur. For this study,

we used the continuous OD. For each lactating cow in a herd, the OD as well as the ELISA plate number was recorded.

Table 1a. Distribution of scores of ordinal risk factors investigated for their relation with the prevalence of paratuberculosis in 97 Danish dairy herds[a].

		Score[b]					
		1	2	3	4	5	Missing
1	Calving and calving area						
a	General facilities (handling, hygiene, space)	12	23	27	15+	2	18
b	Hygiene feeding area	16	48	15+	3+	0	15
c	General hygiene in calving stable	8+	37	35+	4+	0	13
d	General care taking (hoof trimming, feeding, cleanliness)	20	34	30+	2+	0	11
e	Facility maintenance	13	41	29+	3+	0	11
f	Amount of straw in bedding	12	15	17	14+	5	37
g	Hygiene in calving area	5+	18	23	12+	1	38
2	Calves < 2 months						
a	Hygiene in calf pen(s)	21	32	33	10+	0	1
b	Feed hygiene	11	34	33	15+	2	2
c	Animal density	26	33	34+	3+	0	1
d	Care taking of calves	20	39	33+	3+	0	2
3	Calves 2 to 12 months						
a	Hygiene in calf pen(s)	4+	12	45	28+	0	8
b	Feed hygiene	6+	37	37	9+	1	7
c	Animal density	10	29	30	18+	1	9
d	Facility maintenance	10	34	40	6+	1	6
4	Young stock > 12 months						
a	Hygiene in calf pen(s)	5+	8	44	21+	2	17
b	Feed hygiene	8	36	30+	4+	2	17
c	Animal density	7+	20	29	26+	0	15
d	Care taking of calves	13	29	24+	1+	2	28
5	Cows (after 1st calving)						
a	Hygiene in cow stall /area	10	24	26	13+	3	21
b	General hygiene in cow stable	9+	46	37	4+	0	1
c	Feed hygiene	21	44	27	4+	0	1
d	Faecal contamination of cows' skin	17	27	31	21+	0	1
e	Care taking of cows	18	38	37+	3+	0	1
f	Amount of straw in bedding	10	23	25	28+	6	5

[a] +'s between scores show that categories have been combined to obtain reasonable sized classes
[b] The score "1" is the best and "5" the worst for each factor.

Statistical analysis

A modification of a finite mixture model (Gelman et al., 1995) was used, with two mixture components representing infected and non-infected dairy cows, respectively, developed by Nielsen et al. (2007). Previous studies of the milk ELISA (Toft et al., 2005) showed that the log-transformed ODs were reasonably approximated by a Normal distribution. Furthermore, the study demonstrated an effect of age by means of parity and days in milk. The latter association was not linear (on the log scale), therefore, cows were divided into three groups: 0–1 weeks after calving; 2–28 weeks after calving; >28 weeks after calving. Additionally,

cows were grouped into first, second and above second parity to avoid problems with too few observations for older cows. To account for the variation between ELISA plates a random effect of a plate was assumed. Finally, the variance was allowed to vary between the infected and non-infected groups, to reflect that more variation in the log(OD) response was expected in the infected group.

Table 1b. Distribution of scores of dichotomous risk factors investigated for their potential association with the prevalence of paratuberculosis in 97 Danish dairy herds.

		Yes	No	Missing
6	Dry cows			
A	On pasture in summer time	74	17	6
B	Exercise facility	6	77	14
7	Calves < 2 months (after separation from dam)			
A	Calves separated from cows	9	68	20
8	Calves 2 to 12 months			
A	Calves separated from cows	14	58	25
B	On pasture in summer time	65	24	8
9	Young stock > 12 months			
A	Young stock separated from cows	29	63	5
B	On pasture in summer time	85	6	6
10	Cows (after 1st calving)			
A	On pasture in summer time	78	17	2
B	Exercise facility	3	85	9
11				
A	Purchased cows in herd	81	16	0
B	More than 5% purchased cows in herd	60	37	0

Conditioned on the (latent) infection status (*Z*) of the sampled animal, the model assumed that the log-transformed OD values followed a Normal distribution, where the mean and variance were controlled by the infection status and the covariates discussed above. The (latent) infection status (*Z*) was either infected or not infected, i.e. *Z* was Bernoulli distributed, with parameter p defining the probability that the log(OD) was from a truly infected animal. This probability was expected to be influenced by parity and stage of lactation due to the nature of paratuberculosis. Furthermore, it was assumed that there is an effect of herd to allow herd mates to be more similar than non-herd mates. Therefore, the following finite mixture model with two mixture components for the log-transformed OD was specified:

$$\log(OD)_i \mid Z_i \sim N(\mu_i; \sigma^2_{Z_i})$$
$$Z_i \sim Bern(p_i)$$
$$\mu_i = \alpha_{Z_i Parity_i} + \beta_{Z_i DIM_i} + C_{Plate_i}$$
$$C_{Plate} \sim N(0, \sigma^2_C)$$
$$\text{logit}(p_i) = \delta_{Parity_i} + \kappa_{Herd_i} + \psi_{Risk_i}$$
$$\kappa_{Herd} \sim N(0, \sigma^2_\kappa)$$

In this model, log(*OD*)$_i$ is the observed log-transformed OD value for the *i*th sample, which conditional on the Bernoulli distributed infection status (Z_i) follows a Normal distribution with

mean μ_i and variance $\sigma^2_{Z_i}$; $\alpha_{Z_i Parity_i}$ is the combined effect of the (latent) infection status and parity on the mean log(OD), $\beta_{Z_i DIM_i}$ is the combined effect of the (latent) infection status and stage in lactation on the mean log(OD), C_{Plate_i} is the random effect of plate, p_i is the probability that the *i*th sample is from an infected cow, δ_{Parity_i} is the effect of parity on the probability of infection for the *i*th sample, κ_{Herd_i} is the effect of the herd on the probability of infection for the *i*th sample modelled as a random effect of herd (γ_{Herd}), and ψ_{Risk_i} is the effect of a potential risk factor on the probability of infection for the *i*th sample.

Each of the variables from Table 1a, Table 1b and Table 2 were used as risk factors in the model. Only one variable was included in the model at a time, due to the relatively small sample size (at herd level) in this study.

The α's were restricted so that for each parity, the mean log(OD) of the non-infected cows was strictly smaller than the mean log(OD) of the infected cows. This was done to address the implicit unidentifiability of any mixture model (the problem of label switching, i.e. is group 1 the infected or the non-infected?), and to avoid problems with convergence due to the Markov Chain Monte Carlo sampler getting stuck in the trivial solution of all p_i = 1 (or 0).

Parameters were generally given non-informative priors allowing for variation within the relevant range, but restricting extreme values. For example, the priors on α's representing the effect of parity on the non-infected group were given uniform distributions (Uniform(−4,−1)) which allow variation within the relevant range, but prevent extreme values. The α's representing the effect of the infected group, were then specified as the corresponding α for the non-infected, but with a strictly positive contribution added as a Gamma distribution (Gamma(5,3)) with mean (5/3 = 1.67) and standard deviation ($\sqrt{5/9}$ = 0.75) which proved by trial and error to enforce a posterior distribution with different mean log(OD) values for infected and not-infected, while not influencing the estimated difference in any particular way. The prior distribution for all variances were the same, Uniform(0.001,5); the prior distribution for the β's was Uniform(−1,0.5), however, the β's were modelled as an additive effect on α, therefore, most were restricted to 0; finally, the prior for the ψ's was N(0,10).

The odds ratio (OR) for the probability of infection between different levels for a given risk factor was calculated using the first level as a baseline, i.e.

$$OR_j = \exp(\psi_j - \psi_1) \quad \forall : j = 2,\ldots R$$

where R is the number of levels of the risk factor in question.

Statistical inference was made if the 95% credibility posterior interval did not include 1. For factors where the OR was deviating from 1 by more than 0.2 and where this was consistent or deteriorating with a worse score, trends were described to occur.

The model was implemented in WinBUGS (Spiegelhalter et al., 2004), the first 5000 iterations were discarded as a burn-in and the following 10,000 iterations were kept for posterior inference. Convergence of the chain after the initial burn-in was assessed by visual inspection of the time-series plots of selected variables as well as Gelman–Rubin diagnostic plots using three sample chains with different starting values (Brooks and Gelman, 1998).

Table 2. Distribution of herds in various types of housing facilities for various age-groups of cattle, with median posterior odds ratios for paratuberculosis in 97 Danish dairy herds estimated by use of a Bayesian mixture model

	No. of observations	Odds ratio (95% C.P.I.[1])
Calves < 2 months		
Single pen	59	1.0
Other	37	1.08 (0.60, 1.88)
Missing	1	0.007 (0.22, 4.4)
Calves 2 to 12 months		
Single pen or tie stall	12	1.0
Littered common pens	32	3.0 (1.20, 7.2)[§]
Bed stall / slatted floors	18	3.4 (1.04, 9.1)[§]
Combinations	33	2.3 (0.96, 5.8)[#]
Missing	2	0.7 (0.09, 5.6)
Young stock > 12 months		
Tie stall	13	1.0
Littered housing	22	3.7 (1.54, 9.1)[§]
Bed stall / slatted floors	59	2.8 (1.30, 5.8)[§]
Missing	3	0.8 (0.16, 3.4)
Lactating cows		
Tie stall	58	1.0
Bed stall	8	1.0 (0.35, 2.8)
Combinations /other	29	0.85 (0.47, 1.50)
Missing	2	0.81 (0.14, 4.6)
Dry cows		
Tie stall	50	1.0
Littered housing	25	1.0 (0.47, 2.1)
Bed stalls /slatted floors	20	0.34 (0.71, 1.34)
Missing	2	0.18 (0.03, 1.33)
Calving pen(s)		
Used for calving only	11	1.0
Used for calvings and sick cows	52	1.19 (0.56, 2.8)
No specific calving area /pens	34	0.69 (0.30, 1.9)

[1]95% C.P.I.: 95% credibility posterior interval

[#] Indicates a trend, which is not significant at the 5% significance level

[§] Indicates significant results with P<0.05

RESULTS

The distributions of scores for the management factors recorded on an ordinal scale are given in Table 1a. To obtain reasonable sized categories for further analyses, some levels were combined. The distribution of scores for dichotomous management factors are given in Table 1b. In both tables, a number of values are missing. This was particularly for scores on the calving areas, which did not exist separately in 34 herds, and in other instances, some factors could not be assessed. Herds with missing observations for a given factor were included in the analyses as a separate category to assess potential biases of the missing herds.

The type of housing facilities for four age groups, and further subdivided into dry cows, lactating cows and cows in calving areas, are shown in Table 2. Single pens for calves <2 months of age were the dominating housing type in most herds, and tie stalls dominating for

lactating cows. Other types of housing were more evenly distributed for other age-groups. Also shown in Table 2 are the posterior median ORs and the 95% credibility posterior intervals (CPI) for the association between the first and subsequent levels of the risk factors when included as the only herd level risk factor in the model.

Table 3. Median posterior odds ratios (OR) and 95% credibility posterior intervals (95% C.P.I.) for the odds of obtaining paratuberculosis in 97Danish dairy herds estimated by use of a Bayesian mixture model, for ordinal variables describing management factors. OR are given relative to the best score (highest level of hygiene, lowest amount of straw, lowest animal density etc.)[¤]

	$OR_{2nd\ best\ vs.\ best}$	$OR_{3rd\ best\ vs.\ best}$	$OR_{Worst\ vs.\ best}$	$OR_{Missing\ vs.\ best}$
1 Calving and calving area				
General facilities (handling, hygiene, space)	0.95 (0.42, 2.7)	0.73(0.30, 1.77)	0.51 (0.183-1.41)	0.46 (0.173, 1.13)
Hygiene feeding area	1.29 (0.61, 3.3) [#]	2.70 (1.04, 6.8) [§]		1.17 (0.44, 3.2)
General hygiene in calving stable	1.26 (0.71, 2.3)			0.51 (0.23, 1.19)
General care taking (hoof trimming, feeding, cleanliness)	1.07 (0.52, 2.3)	1.39 (0.67, 2.9)		0.51 (0.193, 1.39)
Facility maintenance	1.87 (0.81, 4.2) [#]	1.84 (0.82, 4.4) [#]		0.94 (0.33, 2.5)
Amount of straw in bedding	3.1 (1.27, 10) [#,§]	2.6 (0.94, 7.1) [#,§]	3.0 (1.21, 8.1) [#,§]	1.61 (0.66, 4.0)
Hygiene in calving area	0.97 (0.42, 2.2)	1.11 (0.43, 2.7)		0.83 (0.40, 1.84)
2 Calves < 2 months				
Hygiene in calf pen(s)	0.89 (0.40, 1.9)	0.73 (0.36, 1.48)	0.61 (0.22, 2.2)	0.16 ($7x10^{-3}$-2.5)
Feed hygiene	0.93 (0.42, 2.3)	1.29 (0.56, 3.5)	1.35 (0.54, 3.8)	0.13 ($6x10^{-2}$-1.4)
Animal density	1.28 (0.61, 2.5)	0.78 (0.38, 1.55)		0.22 ($8x10^{-2}$-3.7)
Care taking of calves	0.60 (0.29, 1.20)	0.55 (0.25, 1.09)		0.44 (0.056, 3.1)
3 Calves 2 to 12 months				
Hygiene in calf pen(s)	1.12 (0.43, 2.5)	1.39 (0.56, 3.3)		0.58 (0.128, 2.2)
Feed hygiene	1.06 (0.60, 1.92)	1.11 (0.40, 2.7)		0.38 (0.130, 1.13)
Animal density	1.65 (0.57, 5.0) [#]	1.22 (0.44, 3.7) [#]		1.05 (0.37, 3.8)
Facility maintenance	1.81 (0.63, 4.5) [#]	1.64 (0.55, 4.6) [#]	1.77 (0.46, 7.3) [#]	0.68 (0.160, 2.7)
4 Young stock > 12 months				
Hygiene in calf pen(s)	0.74 (0.30, 2.0)	0.95 (0.28, 2.29)		0.60 (0.21, 1.68)
Feed hygiene	0.83 (0.23, 2.6)	0.74 (0.21, 2.5)		0.71 (0.20, 2.3)
Animal density	2.0 (1.06, 4.4) [§]	2.0 (0.98, 4.4) [#]		1.50 (0.57, 3.5)
Care taking of calves	0.76 (0.30, 1.83)	0.96 (0.43, 2.3)		0.98 (0.40, 2.3)
5 Cows (after 1st calving)				
Hygiene in cow stall /area	0.58 (0.17, 1.83) [#]	0.66 (0.21, 1.85) [#]	0.89 (0.28, 2.7)	0.82 (0.25, 2.8)
General hygiene in cow stable	1.48 (0.89, 2.6) [#]			0.26 (0.011, 4.4)
Feed hygiene	0.73 (0.34, 1.52)	1.44 (0.65, 3.2)		0.20 ($8x10^{-3}$-4.8)
Faecal contamination of cows' skin	1.18 (0.45, 3.2) [#]	1.32 (0.49, 3.1) [#]	1.55 (0.58, 4.3) [#]	0.44 (0.021, 7.0)
Care taking of cows	0.88 (0.43, 2.0)	1.20 (0.55, 2.7)		0.2 (0.007-5.1)
Amount of straw in bedding	1.42 (0.49, 4.5) [#]	1.33 (0.44, 4.0) [#]	1.46 (0.51, 4.4) [#]	1.10 (0.21, 4.9)

[#] Indicates a trend, which is not significant; [§] Indicates significant results with P<0.05; [¤] Best score is score 1, expect for factors 1c, 1g, 2a, 2b, 3a, 3c and 5b, where Scores 1 and 2 were combined as shown in Table 1a.

Littered housing/littered common pens and use of bed stalls/slatted floors among both calves 2–12 months of age and young stock was significant at the 95% confidence level compared to tie stalls.

Median posterior OR and 95% CPI for risk factors recorded on an ordinal scale are given in Table 3. A significant effect of the factor "amount of straw in bedding" in the calving area was seen. The odds of infection were three times higher in herds scoring 2, 3, 4 and 5 relative to score 1 with respect to straw in bedding. The hygiene in the calving cows' feeding areas was also significant, with proper hygiene resulting in the lowest odds of infection. High animal density among young stock resulted in higher odds of infection, with an OR of 2.0 for high and medium density relative to the lowest density of animals.

Table 4. Median posterior odds ratios (OR) and 95% credibility posterior intervals (95% C.P.I.) for the odds of obtaining paratuberculosis in 97 Danish dairy herds estimated by use of a Bayesian mixture model, for dichotomous variables describing management factors.

		$OR_{Yes\ vs\ No}$	$OR_{Yes\ vs.\ Missing}$
6	Dry cows		
A	On pasture in summer time	0.93 (0.42, 2.1)	0.48 (0.146, 1.54)
B	Exercise facility	0.38 (0.131, 1.06)	0.36 (0.102, 1.20)
7	Calves < 2 months (after separation from dam)		
A	Calves separated from cows	0.84 (0.33, 2.3)	0.40 (0.14, 1.40)
8	Calves 2 to 12 months		
A	Calves separated from cows	1.67 (0.84, 3.8) #	1.11 (0.47, 2.6)
B	On pasture in summer time	0.98 (0.53, 2.0)	0.56 (0.22, 1.44)
9	Young stock > 12 months		
A	Young stock separated from cows	1.54 (0.85, 2.8) #	1.12 (0.33, 3.5)
B	On pasture in summer time	0.81 (0.29, 1.92)	0.52 (0.17, 1.56)
10	Cows (after 1st calving)		
A	On pasture in summer time	0.77 (0.38, 1.51)	0.89 (0.16, 5.5)
B	Exercise facility	0.52 (0.11, 2.3)	0.57 (0.093, 2.9)
11			
A	Purchased cows in herd	1.27 (0.65, 2.9)	-
B	More than 5% purchased cows in herd	1.34 (0.74, 2.4)	-

# Indicates a trend, which is not significant

None of the other ordinal factors were significant, but some trends were observed from the size of the OR. Best hygiene in the calving area, among lactating cows and young calves resulted in higher odds of infection than medium and poor levels of hygiene. The following factors all resulted in reduced odds of infection, though they were seen as trends only: high degree of facility maintenance of the calving area and the facilities of calves' 2–12 months of age; the animal density among calves 2–12 months of age; high amount of straw in cows' bedding; low level of faecal contamination of cows' skin. Of the dichotomous management factors, none were significant, but separation of both young stock and calves 2–12 months of age from cows tended to have lower odds of infection (Table 4).

DISCUSSION

Using a Bayesian hierarchical, two-component mixture model, we estimated the effect of a series of potentially important management-related risk factors on the probability of infection with Map. Very few factors were significant at a 95% confidence level: housing type of calves and young stock; hygiene in the feeding area of calving cows; the amount of straw present in the bedding in calving areas; the animal density among young stock. A number of factors yielded ORs, which suggests that the factors were important, but lacking statistical evidence, potentially due to a general high between herd variation.

Housing facilities allowing animals to mix, high animal density, lack of appropriate separation between age-groups and low amounts of straw in bedding are factors for which the analyses suggested that there is an increased risk of infection associated, as it would be expected. Some trends could of course be due to statistical Type I errors, i.e. with 41 factors assessed 1–3 statistically significant management-related risk factors would be expected simply due to random error. In most cases, the ORs are as what could be expected biologically. A statistical Type I error is finding a statistically significant result by chance, without the actual presence of the association between risk factor and disease. In the present study, only factors suspected to be associated with paratuberculosis have been included thereby minimising the risk of Type I errors.

Statistical Type II errors are produced when an actual association between risk factor and disease is not detected. Type II errors were minimised by the use of a mixture-model. However, very few management factors appeared as risk factors by this approach. A potential explanation is the small samples size, with only 97 herds and many missing records. The small sample size at the herd-level is one of the weaker parts of the present study, as can be seen from the many trends, which are biologically sensible but non-significant.

Despite the small sample size, the present study revealed some significant and potentially important findings. Considering the outcome of this study combined with the results of Wells and Wagner (2000) and van Weering et al. (2005), it is questionable as to whether it is actually possible to show which risk factors are more important and which are not, using traditional risk factor studies. A major flaw may be the potential misclassification of recordings of management factors. The management systems may not fit easily into the categories defined on the day of the recording, and may also influence the prevalence more dynamically than can be recorded. Furthermore, risk factor studies for paratuberculosis are troublesome due to a long incubation period with an apparent variable time from infection to infectiousness ranging from few months to many years (Chiodini et al., 1984; van Roermund et al., 2007; Nielsen and Ersbøll, 2006). Therefore, the timing between recording of risk factors and estimates of prevalence can have a major impact on the outcome. In the present study, approximately, 26 months between the two events was chosen because this is the age where the first cows would generally start to sero-convert and shed detectable amounts of Map subsequent to natural infection (Nielsen and Ersbøll, 2006). However, the optimal timing may be closer to 4 years, as most animals sero-convert or become highly infectious in the age-period 3–5 years. Therefore, it is perhaps the management practices and housing of young stock 3–5 years prior to estimation of the prevalence which are important.

An alternative to this kind of ‘time-lagged’ cross-sectional studies and the resulting prevalence data could be to explore cohort-like studies, which would provide incidence data. However, the practical implications and costs of following a large cohort of herds, with

repeated testing of individual animals and follow-up on management procedures prohibits this kind of study for slowly developing infections, such as Map. If the importance of management-related risk factors cannot be assessed by studies of a magnitude such as the current and the ones discussed above, then it is questionable if the effort required is worth it.

The position taken in the newly established Danish control programme against paratuberculosis is to use on-farm risk assessments as learning tools for farmers rather than the use of scores as absolute values, which are comparable across farms. Some management practices may simply be too difficult to accurately characterise and categorise.

CONCLUSION

The use of a multi-level Bayesian mixture model to estimate the probability of infection of individual cows while accounting for age-covariates of individual cows as well as laboratory and herd effects allowed the assessment of herd and management-related risk factors for the within-herd prevalence of Map. In general, the trends of the ORs pointed towards risk factors expected for paratuberculosis. The factors found significant were: animal density among young stock, the amount of straw in the bedding of cows, and the hygiene levels in the feeding areas of cows.

ACKNOWLEDGEMENTS

The Danish Cattle Federation financed this study, partly via Kongeåprojektet and partly via CEPROS-II/7. Nils Toft was financed by a grant from the Danish Research Council.

REFERENCES

Andersen, H.J., Aagaard, K., Skjøth, F., Rattenborg, E., Enevoldsen, C., 2000. Integration of research, development, health promotion, and milk quality assurance in the Danish dairy industry. *In*: Salman, M.D., Morley P., R. Ruch-Gallie (Eds): Proceedings of the 9th symposium of the International Society of Veterinary Epidemiology and Economics, Breckenridge, Colorado, USA, Aug 6-11, pp. 258-260.

Anon., 2004. The risk of transmission of *Mycobacterium avium* subsp. *paratuberculosis* via bovine semen. EFSA Journal 110, 1-59.

Brooks, S.P., Gelman, A. 1998. Alternative methods for monitoring convergence of iterative simulations. J. Comp. Graph. Stat. 7, 434-455.

Chiodini, R.J., van Kruiningen, H.J., Merkal, R.S., 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74, 218-262.

Collins, M. T. 2002. Interpretation of a commercial bovine paratuberculosis enzyme-linked immunosorbent assay by using likelihood ratios. Clin. Diagn. Lab. Immun. 9, 1367–1371.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995, Bayesian Data Analysis, Chapman & Hall /CRC, Boca Raton, Florida, USA.

Groenedaal, H., Nielen, M., Jalvingh, A.W., Horst, S.H., Galligan, D.T., Hesselink, J.W., 2002. A simulation of Johne's disease control. Prev. Vet. Med. 54, 225-245.

Kalis, C.H.J., Hesselink, J.W., Russchen, E.W., Barkema, H.W., Collins, M.T., Visser, I.J.R., 1999. Factors influencing the isolation of *Mycobacterium avium* subsp. *paratuberculosis* from bovine fecal samples. J. Vet. Diagn. Invest. 11, 345-351.

Mortensen, H., Nielsen, S.S., Berg, P., 2004. Genetic variation and heritability of the antibody response to *Mycobacterium avium* subsp. *paratuberculosis* in Danish Holstein cows. J. Dairy Sci. 87, 2108-2113.

Nielsen, S.S. 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subspecies *paratuberculosis* in dairy cattle. J. Vet. Med. B, 49, 384-387.

Nielsen, S.S., Ersbøll, A.K., 2006. Age at occurrence of *Mycobacterium avium* subsp. *paratuberculosis* in naturally infected dairy cows. J. Dairy Sci., 89, 4557-4566.

Nielsen, S.S., Toft, N., 2006. Age-specific characteristics of ELISA and fecal culture for purpose specific testing for paratuberculosis. J. Dairy Sci., 89, 569-579.

Nielsen, S.S., Gröhn, Y.T., Quaas, R.L., Agger, J.F., 2002. Paratuberculosis in dairy cattle: Variation of the antibody-response in offspring attributable to the dam. J, Dairy Sci. 85, 406-412.

Nielsen, S.S., Toft, N., Jørgensen, E., Bibby, B.M., 2007. Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response. Accepted for publication by Prev. Vet. Med.

Spiegelhalter, D.J., Thomas, A., Best, N.G., 2004, WinBUGS Version 1.4 User Manual, MRC Biostatistics Unit, Cambridge, United Kingdom.

Sweeney, R.W., Whitlock, R.H., Rosenberger, A.E., 1992a. *Mycobacterium paratuberculosis* cultured from milk and supramammary lymph nodes of infected asymptomatic cows. J. Clin. Microbiol. 30, 166-171.

Sweeney, R.W., Whitlock, R.H., Rosenberger, A.E., 1992b. *Mycobacterium paratuberculosis* isolated from fetuses of infected cows not manifestin signs of the disease. Am. J. Vet. Res. 53, 477-480.

Taylor, A.W., 1953. Experimental Johne's disease in cattle. J. Comp. Pathol. 63, 355-367.

Toft, N., Nielsen, S.S., Jørgensen, E., 2005. Continuous-data diagnostic tests for paratuberculosis as a multi-stage disease. J. Dairy Sci. 88: 3923-3931.

van Roermund, H. J. W., de Jong, M. C. M., 2006. Horizontal transmission experiment of paratuberculosis. In: Manning, E.J.B., Nielsen, S.S. (Eds.). Proceedings of the 8th International Colloquium on Paratuberculosis, Copenhagen, Denmark, Aug. 14-17, 2005, pp. 672-678.

van Weering, H.J., Mars, M.H., Muskens, J., Middelesch, H., van Schaik, G., 2005. The effect of biosecurity measures for paratuberculosis on the seroprevalence in Dutch dairy herds. In: Manning, E.J.B., Nielsen, S.S. (Eds.). Proceedings of the 8th International Colloquium on Paratuberculosis, Copenhagen, Denmark, Aug. 14-17, 2005, pp. 246-253.

Wells, S.J., Wagner, B.A., 2000. Herd-level risk factors for infection with *Mycobacterium paratuberculosis* in US dairies and association between familiarity of the herd manager with the disease or prior diagnosis of the disease in that herd and use of preventive measures. J. Am. Vet. Med. Assoc. 216, 9, 1450-1457.

Paper XIII

Economy, efficacy and feasibility risk-based control program against paratuberculosis

Kudahl AB, Nielsen SS, Østergaard S, 2008

Published in
Journal of Dairy Science 2008, 91, 4599–4609

Economy, efficacy, and feasibility of a risk-based control program against paratuberculosis

A. B. Kudahl[*1], S. S. Nielsen[†], and S. Østergaard[*]

[*]Faculty of Agricultural Sciences, Dep. of Animal Health, Welfare and Nutrition, University of Aarhus, Blichers Allé 20, P.O. Box 50, DK-8830 Tjele, Denmark,
Tel.: +45 8999 1332, Fax: +45 8999 1525,
[†] Faculty of Life Sciences, Dep. of Large Animal Sciences, University of Copenhagen, Grønnegårdsvej 8, DK-1870 Frederiksberg, Denmark

ABSTRACT

Long-term effects of paratuberculosis on within-herd prevalence and on-farm economy of implementing risk-based control strategies were compared with alternative strategies by using a herd-simulation model. Closing transmission routes is essential for effective control of paratuberculosis. However, many farmers lack the resources to carry out these procedures for all cows in the herd. When using risk-based control strategies 1) all cows are tested quarterly with a milk ELISA, 2) specific cows with a high risk of being infectious are identified, and 3) the farmer can focus only on these infectious animals to close infection routes. In this way the workload can be reduced, making these control strategies more feasible. This study evaluates potential long-term effects of the risk-based approach compared with non-risk-based strategies by simulations conducted with the herd-simulation model PTB-Simherd. Seven control strategies were simulated in herds with initial true herd prevalences of 5, 25, and 50%, respectively. The results predicted the risk-based control strategies to be very efficient and comparable to the best whole-herd strategies in reducing the within-herd prevalence of paratuberculosis with considerably less labor. If infection routes are closed efficiently, prevalence can be reduced to 10% of initial prevalence within 5 to 7 yr. Test-and-cull strategies without closing infection routes were found, by simulation, to be ineffective in reducing prevalence and were not cost-effective methods. The profitability of the various control strategies depends on hourly wages and time spent per cow/calving. Furthermore, simulations show that immediate culling of highly infectious cows is only necessary and cost-effective if infection routes from these cows are not efficiently closed. The risk-based control strategies are recommended in the Danish voluntary control program "Operation Paratuberculosis," which was initiated in February 2006 and now includes 1,220 dairy farmers in Denmark.

INTRODUCTION

Johne's disease or paratuberculosis (PTB) is a chronic infection caused by *Mycobacterium avium* ssp. *paratuberculosis* (MAP), and the infection is widespread in many species, especially in ruminants (Kennedy and Benedictus, 2001). Infections with MAP can cause animal welfare problems and considerable economic losses due to decreased milk production, diarrhea, weight loss, and death in dairy cattle (Chiodini et al., 1984). Furthermore, transmission of MAP through milk, meat, and surface water is suspected to be a potential risk of causing Crohn's disease in humans (Feller et al., 2007). The within-herd prevalence in dairy units has been increasing over the last decades, and several countries have initiated individual control programs, but only with limited success (Kennedy and

Benedictus, 2001). Two of the major challenges in controlling paratuberculosis are that 1) eradication is slow, requiring the farmer to remain motivated for years while diagnosis is hampered by imperfect tests, and 2) effective control implies labor-intensive procedures, often at inconvenient times of the day (e.g., swift removal of the newborn calves) because efficient closure of transmission routes is necessary for control (Groenendaal et al., 2002; Dorshorst et al., 2006; Kudahl et al., 2007). On busy days and in times with shortage of manpower, such work is often given a low priority.

Risk-based strategies can contribute to reduction of both problems mentioned above. They have been implemented in the Danish voluntary control program, Operation Paratuberculosis (OP), initiated in February 2006 by the Danish Cattle Federation. This program consists of 1) frequent testing that substantially increases the probability of identifying infectious animals and 2) focusing special management on only infectious animals when closing infection routes and thereby reducing the amount of labor. The median percentage of animals classified as infectious was 16% among 1,065 herds with updated diagnostic test information by February 22, 2008 (Danish Cattle Federation, Aarhus; unpublished data), which suggest that a significant reduction of animals requiring special management is possible. The strategies are based on the "Bang method," which was used for eradication of bovine tuberculosis in Denmark (Bang, 1908; Bang, 1928). The principle of this method is that all test-positive animals should be managed so that the risk of transmission of the infection is reduced or eliminated (Bang, 1908; Bang, 1928). A prerequisite of the Bang method was frequent testing because the diagnostic test used was imperfect. However, culling of test-positive animals was not a necessity unless the test-positive animals were deemed highly infectious and affected by the infection.

In OP all lactating cows in each herd are tested 4 times per year by using a milk antibody ELISA on milk yield control samples (Nielsen, 2002; Nielsen et al., 2002). The test is optimized to be as sensitive as possible, whereas specificity is of less relevance. Frequent testing enables detection of almost all infected cattle that shed bacteria (Nielsen and Ersbøll, 2006). Cows are grouped into 1) repeatedly ELISA-positive cows, called "red cows" in OP (high-risk cows recommended to be culled); 2) one-time ELISA-positive cows, called "yellow cows" (high-risk cows not recommend to be culled); and 3) ELISA-negative cows, called "green cows" (probably noninfectious cows).

Risk-based management can be established using these risk categorizations (Nielsen et al., 2007). The farmer and herd health advisor need to establish a plan for changes in management and housing systems to reduce transmission of MAP from infectious animals. Transmission of bacteria from potentially infectious cows (with at least 1 positive ELISA) to calves should be closed effectively, and the most infectious cows (repeatedly ELISA-positive) should be culled before next calving (Nielsen et al., 2007). The system allows detection of cows at risk of being infectious, and the farmer only has to optimize management of these cows, thereby saving time and money. However, testing implies a cost along with culling of highly infectious cows. An analysis of the total epidemiological and economic effects of implementing these risk-based strategies compared with non-risk-based strategies in different farms is thereby very relevant and demanded by the farmers. It would, however, be extremely expensive, time-consuming, and practically impossible to evaluate all of these different methodologies using field trials. In addition, MAP infections can have many effects on production parameters, and the effects of different control strategies are very complex.

Therefore, simulation models are needed for estimation of economic and epidemiologic effects.

Various paratuberculosis models have been developed (e.g., Groenendaal et al., 2002; Dorshorst et al., 2006; Kudahl et al., 2007). These models have all been used for the prediction of long-term effects of different control strategies. Most simulation studies agree that management practices interrupting infection routes are essential to control MAP infections in the herd. However, risk-based approaches have not been thoroughly studied.

The present study was based on the PTB-Simherd model, which reflects both direct effects and indirect effects of PTB related to effects on herd dynamics and herd demographics (Kudahl et al., 2007). The aim of our study was to estimate the total epidemiologic, production-related, and economic effects of risk-based strategies recommended in OP compared with different non-risk-based strategies to control MAP infections in dairy herds.

MATERIALS AND METHODS

Simulations were made with the PTB-Simherd model (Kudahl et al., 2007), which is based on Simherd, a dynamic, stochastic, and mechanistic Monte Carlo simulation model simulating a dairy herd including young stock (Østergaard et al., 2003, 2005). The PTB-Simherd model was further developed to simulate the classification system used in OP, where cows are grouped into ELISA-negative, ELISA-oscillating, and repeatedly ELISA-positive groups based on the last 4 milk ELISA results, exactly as practiced in OP. Cows were only classified into the ELISA-negative group if all of the last 4 tests were negative. The repeatedly ELISA-positive group consisted of cows in which at least the last 2 tests were positive. The ELISA-oscillating group consisted of cows with any other combination of positive and negative test results during the last 4 tests. Compared with the sensitivities used in the study of Kudahl et al. (2007), the sensitivities of the milk ELISA were updated based on the study by Nielsen and Toft (2006) (Table 1). Seven scenarios were simulated in a 200-cow dairy herd 10 yr into the future with 500 repetitions. In the PTB-Simherd model both the true infection status and the diagnosed status of each animal is known. Therefore, it is possible to predict the true prevalence of the herd, which is much greater than the apparent (diagnosed) prevalence of the herd. In the results, true prevalences are reported.

Table 1. User-specified values for sensitivities and specificities of milk-ELISA in cows of first and higher parities used in the dairy herd simulation model PTB-Simherd

Infection state	Parity	Sensitivity (%)				Specificity (%)
		7-105 DIM	106-203 DIM	204-300 DIM	301-398 DIM	
Uninfected	1	-	-	-	-	99
Latent infected	1	10	10	10	10	-
Low shedding	1	50	60	65	70	-
High shedding	1	80	80	80	80	-
Clinically affected	1	80	80	80	80	-
Uninfected	>1	-	-	-	-	96
Latent infected	>1	10	10	10	10	-
Low shedding	>1	60	70	75	75	-
High shedding	>1	80	80	80	80	-
Clinically affected	>1	80	80	80	80	-

The 7 scenarios were repeated in 3 herds with initial true prevalences of 5, 25, and 50%, respectively. To compare the effects of risk-based strategies (with different culling strategies) with alternative strategies used in practice, 7 scenarios were specified as follows:

1. No control: A farm with "average" hygiene level; the calf and dam are together for up to 24 h after calving, colostrum is provided from own dam, and the calves are fed with waste milk or bulk tank milk. No testing is performed.
2. Optimized management: All risks of infection are reduced to 5% of risks of strategy 1 (a risk reduction to 0% is assumed to be impossible). In practice this means that all cows are regarded as high-risk cows and PTB management needs to be optimized for all cows and calves in the herd, because no testing is done. Calving areas are thoroughly cleaned before each calving and calves are removed immediately from their dam. Colostrum is used only from own dam if she seems healthy, otherwise from a colostrum bank. Afterward, calves are fed with milk replacer. Generally, a high hygiene level is required, and calves should be raised separately from cows.
3. Optimized management + culling: Like 2, but, in addition, all cows are tested by milk ELISA once a year and positive cows are culled immediately if confirmed with a positive fecal culture. This strategy was included because many farmers want information about the infection status of the cows even though it might not be essential to stop infection routes.
4. Risk-based management 1: Cows are diagnosed by quarterly milk ELISA and grouped into ELISA-negative, ELISA-oscillating, and repeatedly ELISA-positive groups. Management of calves from all cows with at least 1 positive ELISA is performed to close infection routes; that is, ELISA-oscillating cows calve in thoroughly cleaned calving areas that are not used for ELISA-negative cows. Milk (colostrum and other sources of milk) is not used from ELISA-oscillating cows (only from negative cows), and calves are removed from the dam immediately after birth. Calves should be grouped by age. Repeatedly ELISA-positive cows are culled immediately. However, cows with <8 wk to expected calving are allowed to calve because Danish laws prohibit slaughter of cows in late gestation. All calves from these cows are fattened and slaughtered because they have a high risk of being infected with MAP, either before or after calving.
5. Risk-based management 2: As in strategy 4, but repeatedly ELISA-positive cows are not culled until their daily milk yield drops below 10 kg, but not later than 8 wk before next calving. Cows included in the repeatedly ELISA-positive group are not inseminated. This strategy is represented because sometimes farmers want to keep high-yielding ELISA-positive cows until daily milk yield is low.
6. Test and cull based on testing strategies from OP: Immediate culling of repeatedly ELISA-positive cows. No further management of high-risk cows. This strategy was chosen because it already seems to be practiced by some farmers involved in OP, without priorities to closing infection routes. They only use the test results for culling of infectious cows.
7. Test and cull based on testing strategies from OP, but no culling of repeatedly ELISA-positive cows until daily milk yield drops below 10 kg.

All other parameters in the simulation model were specified to represent typical Danish management of a dairy herd of large breeds with capacity for a maximum of 200 cows plus additional young stock. The specified replacement strategy ensures a minimum of 185 cows, and these limits for the herd size define if heifers are sold or purchased. Culling effects are thereby represented directly by the value of the slaughtered cow and the production of the replacement animal. A set of 2006 Danish prices was used to estimate the economic effects. The price of a milk ELISA was set to be US$6, which is the average price in a commercial lab in both Denmark and the United States. The price of a fecal culture test was set to US$32.

The economic costs also include 1 h (hourly wage of US$35) of extra labor per calving for optimizing management as recommended in OP as described for strategy 4. Milk production is the main economic factor affected by MAP infections, but the calculation of the total economic effect also includes the effect on income from sale of heifers, slaughtered animals, and bull calves together with the effects on expenses: feed, purchased heifers, veterinarian, artificial insemination, PTB tests, and additional labor related to control strategies. Future costs and revenues were discounted.

RESULTS

Effect on True Prevalence

Simulation of the testing used in strategies 4 to 7 using 4 annual tests per cows resulted in the grouping of cows described in Table 2. The simulated effects of the 7 control strategies on true prevalence are illustrated in Figures 1, 2, and 3. The ranking of strategies according to effectiveness in reducing MAP prevalence was independent of initial prevalence.

Table 2. Simulated diagnostic test-results compared to the true paratuberculosis infection status of the herd. Based on four quarterly tests with a milk-ELISA, cows are grouped into Repeatedly-ELISA-positive (probably highly infectious), ELISA-oscillating (probably lowly infectious), and ELISA-negative cows (probably non-infectious)

	True infection status				
Simulated diagnostic test-result	Susceptible	Latent	Low shedding	High shedding	Clinical
ELISA-negative	91%	73%	20%	7%	7%
ELISA-oscillating	9%	26%	56%	54%	54%
Repeatedly-ELISA-positive	0%	1%	24%	39%	39%
Total	100%	100%	100%	100%	100%

The curves in Figures 1, 2, and 3 illustrate the importance of closing infection routes by optimizing management. This is shown to be equally effective whether management is optimized for all cows (strategies 2 and 3) or only for high-risk cows (strategies 4 and 5). The simulations also demonstrate that if infection routes are not closed and the tests are only used for culling of repeatedly ELISA-positive cows (strategies 6 and 7), the prevalence is kept more or less constant. The prevalence is not reduced, but the results emphasize the importance of immediate culling of infectious cows rather than keeping them until milk yield is low. However, if infection routes are closed, the time of culling is of no importance in relation to effect on prevalence. If management is carried out as recommended, the prevalence can be reduced to 10% of the initial prevalence within 5 to 7 yr.

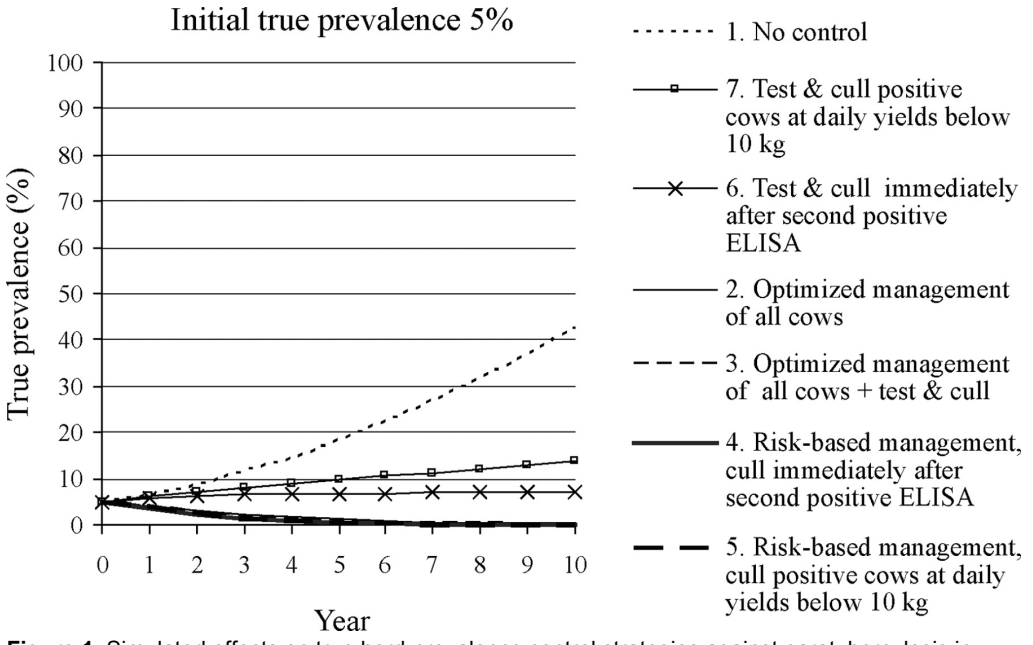


Figure 1. Simulated effects on true herd-prevalence control strategies against paratuberculosis in dairy herds with an initial true herd prevalence of 5%. Risk-based strategies (strategies 4 and 5) were compared with no control (strategy 1) and non-risk-based strategies (2, 3, 6, and 7).

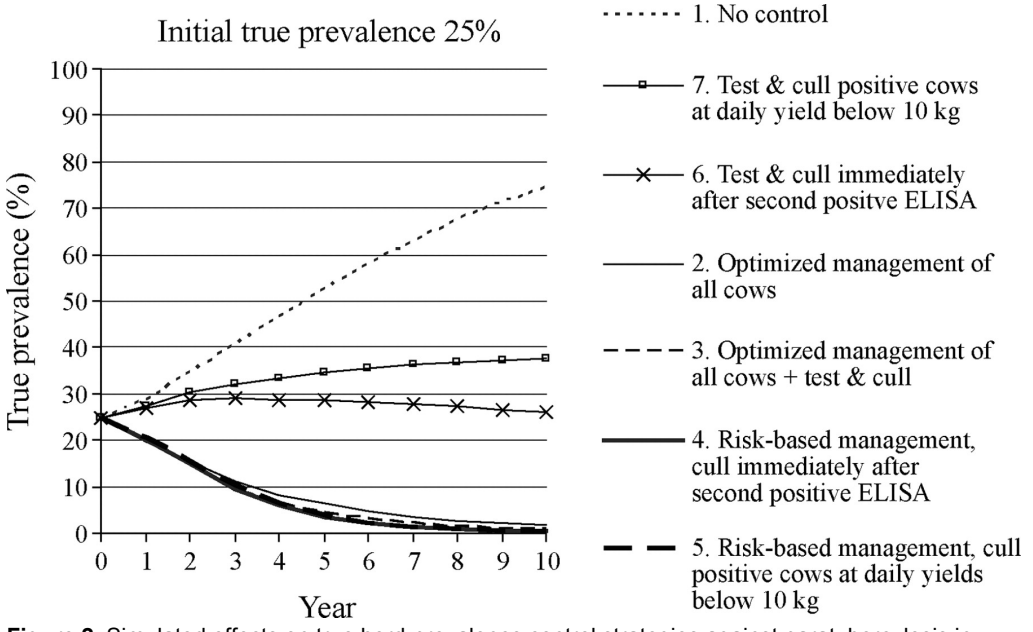


Figure 2. Simulated effects on true herd-prevalence control strategies against paratuberculosis in dairy herds with an initial true herd prevalence of 25%. Risk-based strategies (strategies 4 and 5) were compared with no control (strategy 1) and non-risk-based strategies (2, 3, 6, and 7).

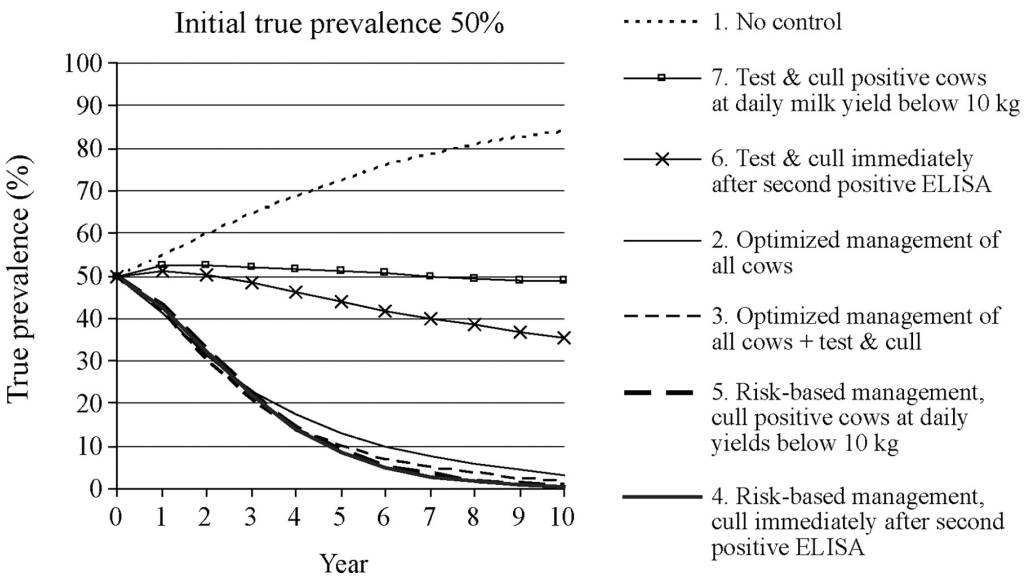


Figure 3. Simulated effects on true herd-prevalence control strategies against paratuberculosis in dairy herds with an initial true herd prevalence of 50%. Risk-based strategies (strategies 4 and 5) were compared with no control (strategy 1) and non-risk-based strategies (2, 3, 6, and 7).

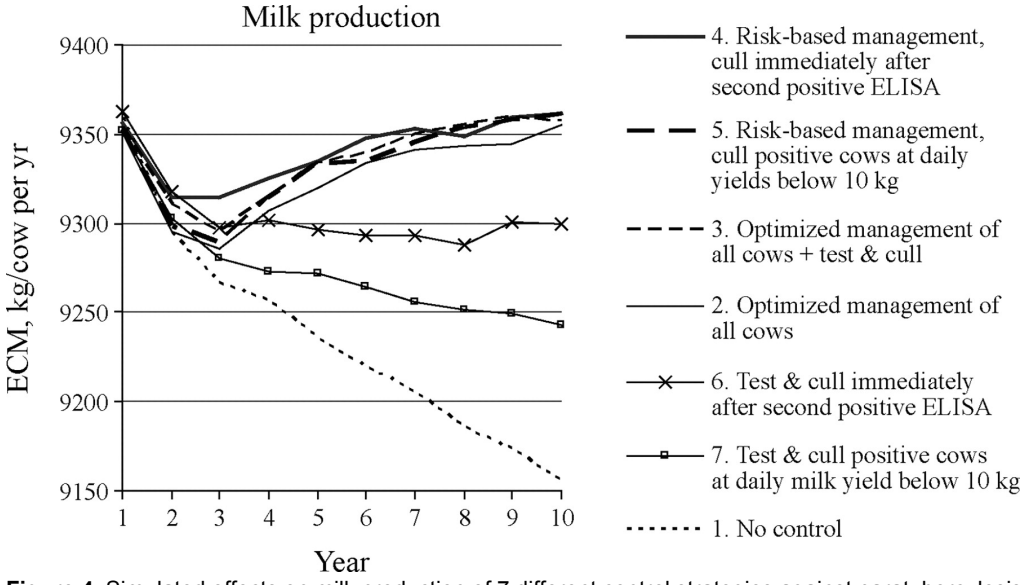


Figure 4. Simulated effects on milk production of 7 different control strategies against paratuberculosis in a dairy herd with initial true prevalence of 25%. Risk-based strategies (strategies 4 and 5) are compared with no control (strategy 1) and non-risk-based strategies (2, 3, 6, and 7).

Economic Effects

The effect on milk production of the different control strategies on a farm with initial true prevalence of 25% is shown in Figure 4 as an example of the main economic factor. Independent of control strategy, it takes 2 to 3 yr before milk production stops decreasing and starts to increase. If no control is performed, milk yield will decrease continuously.

The total economic effect on net annual revenue per cow per year is illustrated in Figures 5, 6, and 7. The economic results of the first 2 yr can be slightly misleading as they are strongly affected by the status of the initial herd. Although true prevalence is still less than 20% (Figure 5, yr 1 to 6) none of the control strategies seem cost effective. However, if no control strategies are implemented after this point (strategy 1), the simulations predict that the losses will increase constantly with increasing prevalence (Figures 6 and 7). If a control strategy is implemented when true prevalence is 25%, costs of all strategies exceed the profit from reducing prevalence in the first 3 to 4 yr, and it will thus be tempting to let things slide. But over a period of 10 yr and with an initial true prevalence above 25%, the risk-based strategies (strategies 4 and 5) seem to be the most favorable economically for control of MAP (Figures 6 and 7). Culling of repeatedly ELISA-positive cows immediately (strategy 4) instead of waiting until their daily milk yield is <10 kg (strategy 5) is of no economic importance as long as infection routes are closed. If routes are not closed, it is important to cull cows as soon as they have their second consecutive positive ELISA (strategies 6 and 7). The best alternative strategy compared with risk-based management is to carry out optimal management on all cows without testing and culling (strategy 2). Supplementing this strategy with testing and culling (strategy 3) is an extra cost that is never recouped and does not make the strategy more effective.

These conclusions are, however, based on the assumption that optimizing management with the aim of closing infection routes takes 1 h (hourly wage of US$35) per calving/calf in total (Figure 8). Herds are, however, different and in some modern herds it may be easier to change management. Figure 9 shows the economic effects in a herd in which the extra management takes only 20 min per calf. In this case, strategy 2 (optimizing management of all cows without testing) is more profitable than the risk-based strategies (4 and 5) the first 3 yr because there are no expenses for tests and culling. From this point, that strategy (with or without using tests once a year) has the same economic effect as risk-based management. This means that if extra labor to optimize management is estimated to take <20 min per calving/calf (or cost <US$12) it is economically favorable to optimize management of all cows (strategy 2) without testing and culling (strategy 3) instead of using risk-based management (strategies 4 and 5), which implies costs for tests and culling.

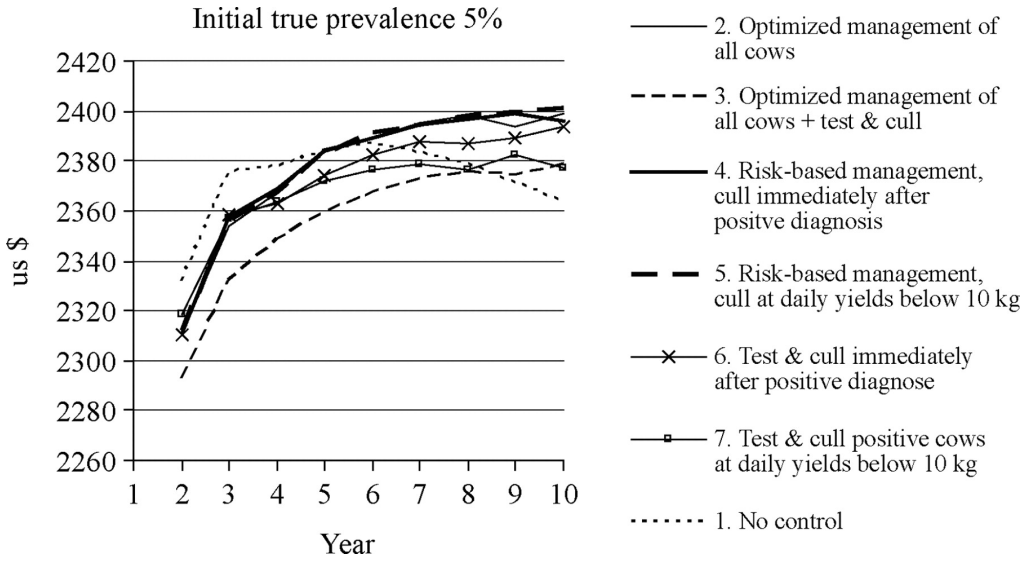


Figure 5. Simulated economic effects (calculated as net annual revenue/cow per year, US$) of control strategies against paratuberculosis in dairy herds with an initial true herd prevalence of 5%. Risk-based strategies (strategies 4 and 5) are compared with no control (strategy 1) and non-risk-based strategies (2, 3, 6, and 7).

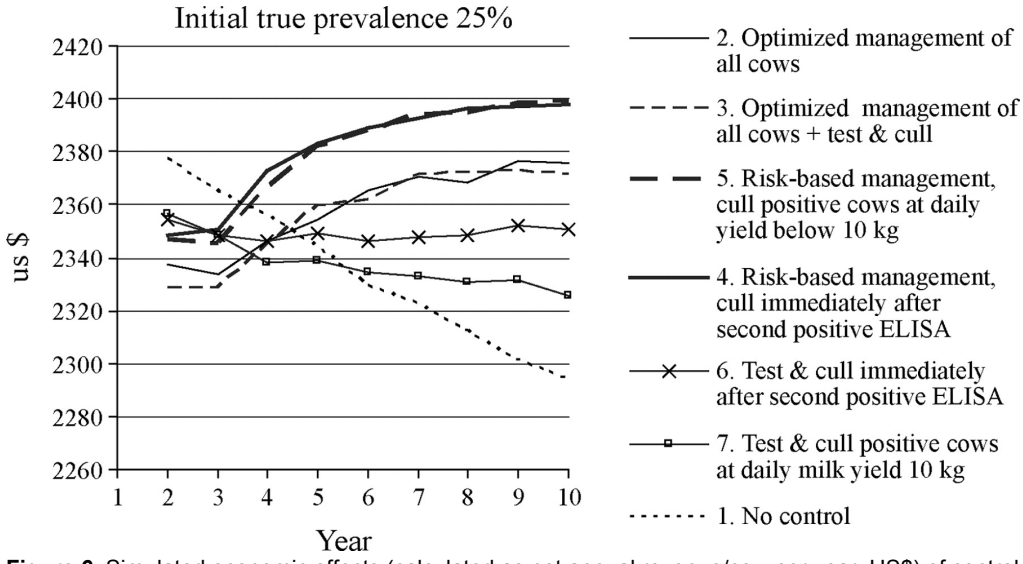


Figure 6. Simulated economic effects (calculated as net annual revenue/cow per year, US$) of control strategies against paratuberculosis in dairy herds with an initial true herd prevalence of 25%. Risk-based strategies (strategies 4 and 5) are compared with no control (strategy 1) and non-risk-based strategies (2, 3, 6, and 7).

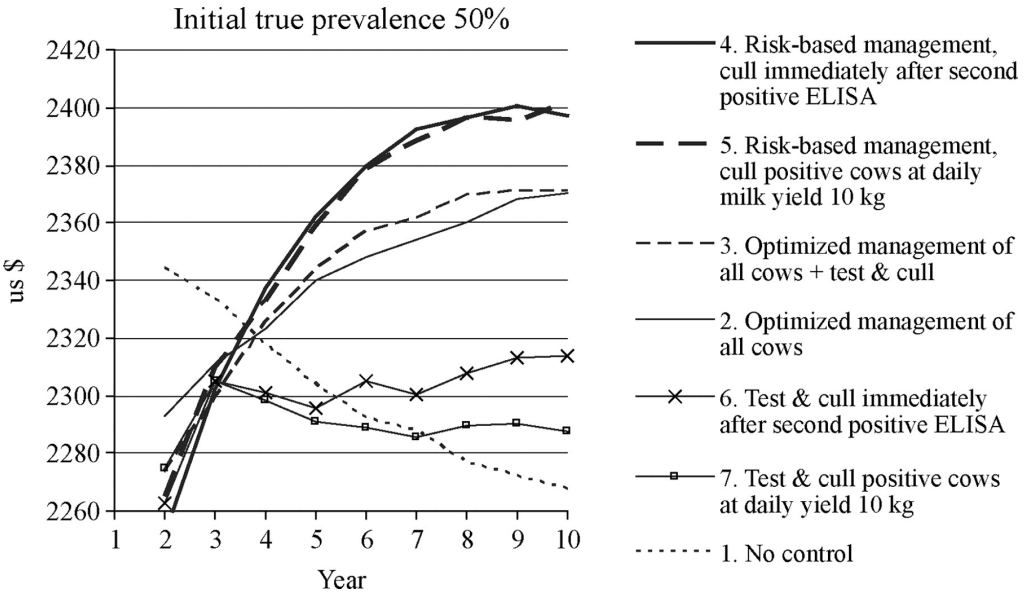


Figure 7. Simulated economic effects (calculated as net annual revenue/cow per year, US$) of control strategies against paratuberculosis in dairy herds with an initial true herd prevalence of 50%. Risk-based strategies (strategies 4 and 5) are compared with no control (strategy 1) and non-risk-based strategies (2, 3, 6, and 7).

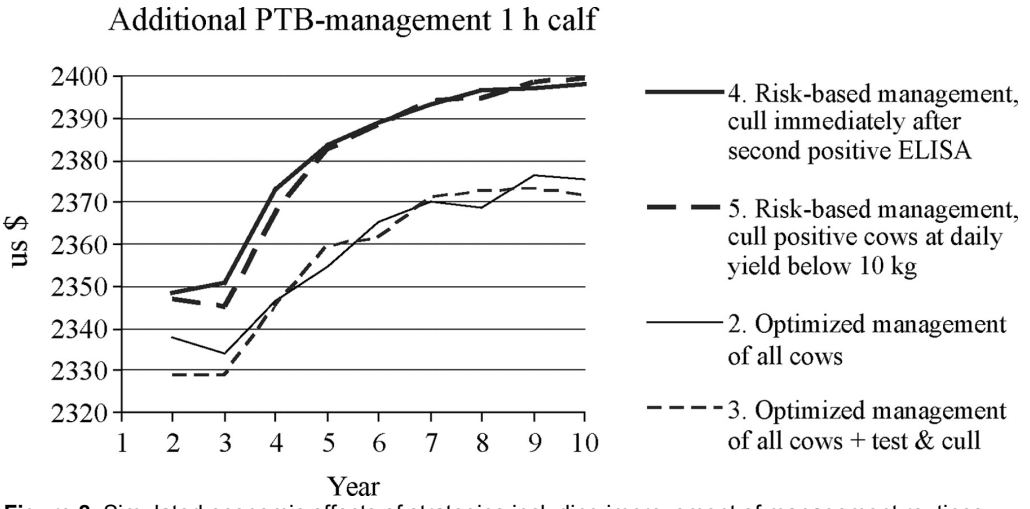


Figure 8. Simulated economic effects of strategies including improvement of management routines where this improvement takes additional 1 h per calving/calf. Thick lines illustrate Operation Paratuberculosis with special management of only infectious cows. Thin lines are strategies in which all cows of the herd are managed to close infection routes using no tests.

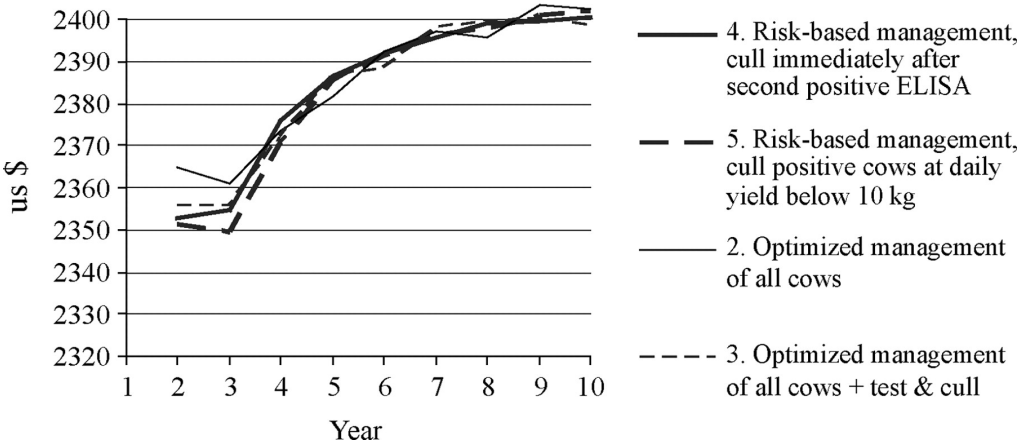


Figure 9. Simulated economic effects of strategies including improvement of management routines in which this improvement takes only 20 min per calving/calf. Thick lines illustrate Operation Paratuberculosis with special management of only infectious cows. Thin lines are strategies in which all cows of the herd are managed to close infection routes using no tests.

DISCUSSION

Our simulations showed that risk-based approaches to control MAP infections are more cost effective than non-risk-based approaches, although the results are highly dependent on the time spent per calving and the hourly wages for management of infectious cows. The resulting simulations predicted risk-based management to be very effective in reducing the prevalence of MAP infection, independent of initial true herd prevalence. Closing infection routes by special management was, however, essential for reducing prevalence effectively, irrespective of the strategy chosen. The importance of closing infection routes was also emphasized by Bang (1908, 1928) and it was predicted by other simulation studies (Groenendaal et al., 2002; Dorshorst et al., 2006, Kudahl et al., 2007). Test and cull strategies were not cost effective in any of the herds simulated in this study. The results show that, for eradication purposes, the only effective strategies are risk-based management (strategies 4 and 5) and those implying optimized management of the whole herd (strategies 2 and 3). When comparing these strategies, cost effectiveness depends directly on time (wages) and the amount of time necessary for optimizing management for each calf: risk-based management is profitable whenever this special management exceeded US$12 per calving (under Danish conditions).

Available antibody ELISA tests have large variation in their ability to detect infectious animals (Nielsen and Toft, 2008). Moreover, the cut-off values of the individual tests regulate the sensitivity. Choosing a low cut-off value results in high sensitivity and low specificity; the opposite occurs with a high cut-off value. The cut-off value can thereby be adjusted to meet the aims of the control strategy. The cut-off value for the milk ELISA used in our strategies was chosen to obtain a high sensitivity of 50 to 80% for shedding cows (depending on infection stage and parity) and 10% for latently infected cows. Thereby, specificity was reduced to 96% for second-parity and higher parity cows. Other ELISA tests are available on the market, and therefore the simulations were repeated with 2 other levels of sensitivity: 1) for highly infected and clinically ill animals, sensitivity was reduced to 60% and specificity increased to 97%, and 2) for all infectious animals, sensitivity was reduced to 40% and

specificity increased to 99.9%. Compared with the high-sensitivity ELISA, the results were that the process of reducing true herd prevalence from 25 to 5% was delayed by 6 mo and 3 yr, respectively. The net annual revenue per year-cow was reduced by US$3 to 5 with a sensitivity of 60%, whereas a sensitivity of 40% actually increased the income in the first 3 yr because fewer animals were culled. However, after 3 yr the net annual revenue was reduced by US$5 to 10 per year-cow.

Another sensitivity analysis was made on the testing intervals. In our strategy, we suggested that all cows were tested quarterly. Testing 2 or 6 times per year was also simulated. By testing 6 times per year a reduction of herd prevalence from 25 to 5% was obtained 4 to 5 mo earlier than when testing quarterly. By testing twice a year, the reduction was delayed by 2 yr. Over a period of 10 yr, there is no economic difference between testing twice a year or quarterly, whereas testing 6 times a year costs an additional US$3 to 5 per cow per year. When considering the effects on both prevalence and economy, we conclude that the optimal test schedule is quarterly.

Earlier studies have shown that herds with poor reproduction and high PTB prevalence react differently to control strategies (Kudahl et al. 2007). Simulations of the 7 control strategies in a herd with a heat detection rate of 40% and an initial prevalence of 50% showed that the culling of repeatedly ELISA-positive cows in such a high-prevalence herd with poor reproduction is very costly because of the scarcity of replacement animals. This made management strategies without testing and culling much more attractive and profitable whenever costs for management were <US$63 per calving. Only above this limit was it attractive to use risk-based strategies and to initiate the culling of repeatedly ELISA-positive cows.

Dorshorst et al. (2006) also suggested control strategies that include special management of infectious groups of animals based on different testing systems and management levels. But they found that testing was profitable under any conditions as long as it was used for managing test-positive cows. They have estimated the total costs for optimizing management to be US$260.50/cow per year, which is about 8 times our estimated cost. The high cost makes detection of risk animals by testing much more profitable. The simulated farming systems are, however, very different from the systems simulated in our study. Dorshorst et al. (2006) included one extra full-time employee to a 100-cow herd for optimizing management and investments in new calf hutches. In our studies, we simulated a typical Danish farming system with a new stable (including calf hutches), a high degree of automation, and scarcity of manpower. Under these conditions, costs for the change in management routines to close infection routes are much lower.

In our economic results, all variable costs and income related to the dairy herd and its production are included in the calculations, together with the additional costs for labor when optimizing management. However, there may be farms with other economic effects of controlling MAP infection that we have not included here: investments in additional calving areas, electronic surveillance of calving areas, cleaning equipment, and additional freezing capacity for a colostrum bank. On the other hand there may also be some positive side effects: 1) closing infection routes for MAP will also close infection routes for several other fecal-orally transmitted infections among the calves, and calf mortality may decrease, body weight gain may increase, and over time there will be more and better replacement animals in the herd; and 2) the sales value of the herd will increase as the prevalence of paratuberculosis goes down.

The importance of immediate culling of repeatedly ELISA-positive cows has been questioned by the farmers, particularly if there is no apparent effect on milk yield. According to the simulations immediate culling only seems important if infection routes are not closed. However, the results also show that by postponing the time of slaughter, there is only a minor economic benefit from keeping these cows to the end of lactation (particularly when infection routes are closed), because the majority already exhibit reduced milk production, and some cows lose weight before slaughter. In practice, such a strategy would possibly be very selective, and only high-yielding ELISA-positive cows would be allowed to finish lactation, which is more profitable than keeping all ELISA-positive cows.

Even if a cow is diagnosed as repeatedly ELISA-positive in late pregnancy, Danish law does not allow her to be culled until after calving. It is recommended that her calf is separated from other heifer calves, fattened, and culled. There is a high risk of transplacental infection and probably also a greater risk of shedding bacteria at an early age. This risk of calf-to-calf infection is not included in the simulation model. In scenarios without removal of calves from high-risk cows, the transmission of MAP in the herd is likely underestimated. In practice, however, risk-based strategies recommend that these calves are raised and sold together with bull calves; therefore, the risk of horizontal transmission between heifer-calves is minimized.

In general, the simulations predict the costs of all control strategies against PTB to exceed the increased income (from increased milk production) in the first 3 to 7 yr depending on strategy. Alternatively, the result of not controlling PTB over a longer time span is constantly increasing production losses, and thereby constantly decreasing net annual revenue. This reinforces the point that control of paratuberculosis requires patience and persistence from the farmers.

Operation Paratuberculosis was initiated in February 2006 and there are still only a few results available, and analysis of the effects of using the strategies recommended have not previously been performed. Validation of the simulations is therefore only possible by sensitivity analysis and face validation (Sørensen, 1990). Face validity is obtained by asking people who have insight into the system whether the conceptual model is reasonable by means of flowcharts (Sargent, 1982). One technique in face validation is loop analysis, in which all feedbacks in the model are examined for polarity, gain, and delay. Previous studies have confirmed the importance of closing transmission routes and found poor effects of controlling PTB only by testing and culling (Groenendaal et al., 2002; Dorshorst et al., 2006; Kudahl et al., 2007).

Although these simulations indicate that, under some conditions it is more profitable to use no tests and manage all cows optimally instead of joining OP, experiences from several countries show that this strategy is only practicable for a few farmers (Kennedy and Benedictus, 2001; Groenendaal et al., 2003). The burden of this labor is considered heavy, often occurring at inconvenient hours. Sources for the extra hours of labor are often not available because of the scarcity of manpower. In these cases, the strategies of OP are probably more manageable for effective control of PTB. Furthermore, regular tests can be of important informative and motivational value in control programs because they reflect whether there is an effect on prevalence.

Based on the large economic effects of cullings on farms with poor reproduction, the importance of culling ELISA-positive cows on these farms could be reconsidered and

analyzed by further simulations. However, there remain important ethical and biosecurity reasons to quickly cull the most infectious and clinically diseased cows.

Another aspect not included in the simulation model is the communicative learning and motivating process that can be achieved by risk assessments and use of test results for monitoring in collaboration between the farmer and the advisor. The economic value of this process is not included, but the educating and motivating effect on the farmer is considered to be crucial to the success of the control program.

REFERENCES

Bang B. 1908. The Bang Method for the repression of tuberculosis in cattle. Commonwealth of Pennsylvania, Department of Agriculture, Harrisburg, Pennsylvania, USA. Bulletin no. 172.

Bang B. 1928. Tuberculosis in cattle. J. Am. Vet. Med. Assoc. 72: 20-25.

Chiodini R ., HJ van Kruiningen, RS Merkal. 1984. Ruminant paratuberculosis (Johne's disease): the current status and future prospects. Cornell Vet. 74:218-262.

Dorshorst NC, MT Collins JE Lombard. 2006. Decision analysis model for paratuberculosis control in commercial dairy herds. Prev. Vet. Med. 75:92-122.

Feller M, K Huwiler, R Stephan, E Altpeter, A Shang, H Furrer, GE Pfyffer, T Jemmi, A Baumgartner, M Egger. 2007. *Mycobacterium avium* subspecies *paratuberculosis* and Crohn's disease: a systematic review and meta-analysis. Lancet Infect. Dis. 7:6007-613.

Groenendaal H, M Nielen, AW Jalvingh, SH Horst, D T Galligan,JW Hesselink. 2002. A Simulation of Johne's disease control, Prev. Vet. Med. 54:225-245.

Groenendaal H, M Nielen, JW Hesselink. 2003. Development of the Dutch Johne's disease control program supported by a simulation model. Prev. Vet. Med. 60:69-90

Kennedy DJ, G Benedictus 2001. Control of *Mycobacterium avium* subsp. *paratuberculosis* infection in agricultural species. Rev. sci. tech. Off. Int. Epizoot. 20:151–179.

Kudahl AB, S Østergaard, JT Sørensen, SS Nielsen. 2007. A stochastic model simulating paratuberculosis in a dairy herd. Prev. Vet. Med. 78:97-117.

Nielsen SS 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to *Mycobacterium avium* subsp. *Paratuberculosis*. J. Vet. Med. B 49:384-387.

Nielsen SS, AK Ersbøll. 2006. Age at occurrence of *Mycobacterium avium* subsp. *paratuberulosis* in naturally infected dairy cows. J. Dairy Sci. 89:4557-4566.

Nielsen SS, C Grønbæk JF Agger, H Houe. 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53:191-204.

Nielsen SS, ØR Jepsen, K Aagaard. 2007. Control programme for paratuberculosis in Denmark. Proceedings of the 1st ParaTB Forum. Bulletin of the International Dairy Federation 410: 23-29.

Nielsen SS, N Toft. 2006. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. J. Dairy Sci. 89:569-79.

Nielsen SS, Toft N. 2008. Antemortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques. Vet. Microb. 129, 3-4, pp 217-235.

Østergaard S, MGG Chagunda, NC Friggens, TW Bennedsgaard, IC. Klaas. 2005. A stochastic model simulating pathogen-specific mastitis control in a dairy herd. J. Dairy Sci. 88:4243–4257.

Østergaard S, JT. Sørensen, H Houe. 2003. A stochastic model simulating milk fever in a dairy herd. Prev. Vet. Med. 58:125–143.

Sargent RG. 1982. Verification and validation of simulation models. Pages 159-169 in: Progress in modelling and Simulation. F.E. Cellier (Editor) Academic Press, London.

Sørensen JT, 1990. Validation of livestock herd simulation models: a review. Livest. Prod. Sci. 26:79-90.